DETECTING RNA FOLDING AND DYNAMICS OF THE TWISTER RIBOZYME AND THE 16S RIBOSOMAL RNA

by
Arthur Korman

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
October 2019

Abstract

Ribozymes are ribonucleic acid (RNA) enzymes capable of catalyzing chemical reactions. Their ability to catalyze reactions is dependent on the formation of proper secondary and tertiary folds. RNA folding begins as the nascent RNA is being synthesized. Structures that are formed early by the 5′ RNA will rearrange, allowing for more mature folds to take hold once the full-length RNA is synthesized. The final fold is the result of constant shifts in RNA structure as the RNA travels through its rugged folding pathway, until the full RNA has been synthesized and the RNA has reached its thermodynamically stable structure. In this dissertation, the folding dynamics of two RNAs, the bacterial 16S ribosomal RNA (rRNA) and the twister ribozyme, are studied. The 16S rRNA undergoes many conformational changes during 30S ribosome assembly, including a critical structural rearrangement during late stage of 30S subunit biogenesis between two competing structures, helix 1 and a hypothesized alternative helix 1. Using biochemical and analytical techniques, helix switching that occurs between the competing helical structures during 30S subunit reconstitution is examined. Unlike the large 16S rRNA, the Twister ribozyme is a small RNA that undergoes dynamic structural changes required for self-scission. The folding dynamics of twister RNA were measured in real time using single molecule Förster resonance energy transfer (smFRET). A photo-activatable chemical group, called a photocage, was able to trap the twister RNA in its secondary fold. Tertiary folding was temporally controlled by activating the photocage with a 405 nm laser. By controlling when secondary and tertiary folding occurred, folding behavior before and after tertiary interactions formed and self-scission occurred can be monitored directly. Studying this dynamic behavior has provided insight into the rate of folding and how environmental conditions influence its folding dynamics.

Understanding RNA dynamics is critical for understanding RNA function and creates a future opportunity for predicting RNA folding and function from its sequence.

Advisory Committee:

Dr. Sarah A. Woodson

Dr. Gregory D. Bowman

Dr. Christian Kaiser

Dr. Stephen D. Fried

Dedication

To Neha, thank you for believing in me and pushing me to be better

Acknowledgements

I am thankful to Dr. Sarah A. Woodson for allowing me to be in her lab, and for giving me an opportunity to do an internship during my time here. Thank you to the Greenberg and Ha labs for their help. I am grateful to all of my committee members for their advice, support, and for donating their time to help me graduate.

I got a chance to work with some amazing people in the Woodson lab, thanks to everyone for guiding me and teaching me as I went along. Subrata, Indra, and Maggie thank you for your post-doctoral wisdom during coffee hours, I would not have been able to survive without it.

I am thankful to my family for bringing me to the United States and giving me opportunities which they never had. Their sacrifice led to my schooling, which has set me on a path which no one in my family has ever taken. I will never be able to repay them. A big thank you to my grandfather, Yakov Rabinovich, for teaching me my multiplication tables and supporting my education.

A big thank you to Dr. Wali A. Karzai for being an amazing mentor and a guide in my adult life. I will forever be in his debt for allowing me to study in his lab and learn the scientific process from him. Thank you for giving me the opportunity to get involved in biochemistry and molecular biology, for giving me a job right after graduation, and for generally dealing with me. You were right, sorry for not listening.

A great, big, huge thank you to my wife Neha, I would be lost without you. Thank you for putting up with me and supporting me. Looking forward to the future, I don't know what I want to do, but I know I don't want to be a graduate student.

Table of Contents

List of Tables

List of Figures

Chapter 1: RNA Folding

Introduction

RNA is a versatile molecule capable of storing information as well as performing catalytic functions required for survival. This versatility has made it a prime candidate for being the first polymer to arise in an abiotic environment, leading to the RNA world hypothesis (reviewed in (Higgs and Lehman 2015)). This hypothesis relies on the idea that RNA or RNA-like molecules arose at the beginning of life that were capable of storing information and catalyzing self- replication. Today, RNA continues to perform catalysis but is relied upon to a smaller degree. DNA has evolved as a more stable means of storing information, while proteins are capable of rapid, controllable, catalysis. Messenger RNAs (mRNAs) are used to store and transfer information, while ribonucleic acids enzymes or ribozymes are used to perform peptidyl transfer reaction in the ribosome, splicing of mRNAs, and signal transduction via metabolite sensing and self-cleavage.

Information carried by RNA is within its primary sequence, the nucleotide sequence that is read by the ribosome while synthesizes proteins, or by reverse transcriptase while synthesize complimentary DNA (cDNA). However, the catalytic activity of RNA arises only when it takes on a proper secondary and tertiary structure. This folding happens in stages and is done as soon as the RNA starts to be synthesized by RNA polymerase or co-transcriptionally (reviewed in (Schroeder et al. 2002; Ganser et al. 2019)). The folding of RNA is a complex process, with its local secondary structural folding occurring in parallel with its tertiary fold, which arises from long range interactions (Schroeder, Barta, and Semrad 2004). Certain secondary and tertiary structures that fold initially during transcription will refold to bring about structures with lower

free energy during the RNAs journey through the folding landscape. Secondary folds are relatively stable and form due to Watson-Crick base pairing resulting in the formation of helices and loops. Most RNAs are treated as unfolded when they only contain secondary structures because these structures are so stable, and its rarely observed how an unstructured and unfolded RNA develops its secondary fold (reviewed in (Draper, Grilley, and Soto 2005)). Tertiary structures occur from distant RNA sequences coming closer together to stabilize a specific structure, resulting in the stabilization of an active site responsible for performing catalysis. Because the RNA has a negatively charged backbone, it is difficult to bring sequences together in a tertiary fold without the help of positively charged metal ions, proteins or other positively charged cellular metabolites such as spermine and spermidine.

Metal ions stabilize tertiary folds

K^+ and Mg^{2+} are the two most abundant ions within the cell (Romani 2011). Both of these ions have a role to play in stabilizing RNA structure by preventing the negative backbone from repulsing itself through their positive charge. Mg^{2+} is the smaller of the two ions and has the larger positive charge, making it great at stabilizing RNA structures. Previously it was thought that the metal ions worked by binding to specific sites within the RNA, stabilizing the binding site via induced fit, and allowing for the RNA to fold. Now it is much more accepted that metal ions play a general role in folding, by interacting with the RNA through electrostatic forces, and can influence both RNA folding and other ion binding long distances from the point at which they are located (Draper, Grilley, and Soto 2005). Increasing Mg^{2+} concentration has been documented to collapse the RNA structure into a condensed form that is not necessarily its native

fold (reviewed in (Woodson 2010)). However, by allowing RNA to form its condensed structure it might allow for the active site to refold, eventually ending up at the native, active structure.

The presence of Mg^{2+} is important for *in vitro* ribosomal 30S subunit reconstitution, as well for the *in vitro* folding and activity of most ribozymes. Twister ribozymes, a bioinformatically identified self-cleaving ribozyme family, are capable of folding in Mg^{2+} but have a lower midpoint of cleavage in the presence of other divalent metal ions such as nickel, cobalt and manganese, suggesting a preference for other metal ions (Roth et al. 2014; Panja et al. 2017a). In order to understand the folding dynamics of twister under different metal ion concentrations, and to understand the effect these ions have on folding and cleavage, I attached a fast activating photocage to the ribozyme. Using single molecule fluorescence resonance energy transfer (smFRET) spectroscopy, I studied the dynamics of twister ribozyme from the very start of its ability to fold from secondary to tertiary structure, in the presence of varied concentrations of metal ions.

Proteins can play a role in RNA folding

Unlike metal ions which work through passive means, proteins are capable of folding RNA in a more active manner (Schroeder et al. 2002). Although the functions of the ribosomes are performed by the RNA, proteins play a role in helping the RNA fold by either stabilizing or destabilizing specific folds and interactions (Weeks 1997). Most ribosomal proteins bind to ribosomal RNA and restructure the RNA and themselves. The initial protein-RNA or RNA-RNA interaction changes over the lifetime of the RNA-protein complex until it arrives at its native fold. While ribosomal proteins remain bound to the ribosomal RNA, other proteins, such as ribosomal assembly factors, are capable of binding, guiding folding, checking for proper RNA

dynamics and then dissociating from the complex. Proteins help guide the RNA through its folding landscape, allowing for quick folding of the RNA by preventing the stabilization of unproductive folded states.

Defining the Ribosome

The ribosome is a ribonucleoprotein complex that is responsible for the translation of messenger RNA (mRNA) to proteins. It is present in all domains of life, and the core of the ribosome is well-conserved from bacteria to humans (Melnikov et al. 2012). It is composed of two subunits, a large subunit and a small subunit, with each subunit itself being composed of both RNA and proteins (Noller 1991; Yusupov et al. 2001). The bacterial 70S ribosome is composed of small subunit (30S) and a large subunit (50S), all named after their sedimentation coefficients, which are measured in units of Svedberg (S) and describes their ability to pellet during ultracentrifugation. The *E.coli* 30S subunit is composed of a 1542 nucleotide (nt) ribosomal RNA (rRNA) and 20-21 ribosomal proteins (r-proteins) (Shajani, Sykes, and Williamson 2011). The 50S subunit is composed of two rRNAs, a 2,904 nt 23S, and a 120 nt 5S, with a total of 33 proteins (Moore and Steitz 2002). Each subunit plays a role in translating mRNA into proteins, and its proper function can only come about when the rRNA is properly folded.

Within the 70S ribosome, the large subunit (LSU) contains the peptidyl transferase center, which is responsible for the synthesis of the nascent protein, making the ribosome a *bona fide* ribozyme (Noller, Hoffarth, and Zimniak 1992; Nissen et al. 2000). The small subunit (SSU) is responsible for binding mRNA, translocating, and accurately decoding it (Noller 1991; Carter et al. 2000). Proper folding of the rRNA is essential for all of these functions and the cell cannot survive without making sure that rRNA is folded, with most antibiotics targeting the ribosomes

ability to translate RNAs into proteins. Ribosomal proteins coat the rRNA, guide its folding, and support the binding of elongation factors during translation (Noller 2017). If the rRNA is misfolded, or a r-protein is under the stoichiometric amount within the cell, this can result in ribosomopathies.

The *rrn* operons are highly regulated

E.coli contain seven operons which code for the 16S rRNA, 23S rRNA and the 5S rRNA, with interspersed tRNAs (Sarmientos et al. 1983; Dahlberg 1999). Based on the environmental conditions, the total number of ribosomes present in a cell can vary between 20,000 and 70,000 (Condon et al. 1999). This means that the cell has to transcribe, mature, maintain and then degrade thousands of ribosome complexes depending on the needs for its survival. The cell also maintains the levels of the rRNA, even in the presence of plasmid DNA containing different *rrn* operons (Brosius et al. 1981; Gourse, Boer, and S 1986). The concentration of r-proteins is also highly regulated, with many r-proteins participating in negative feedback loops to control their own synthesis (Razi and Ortega 2017; Torres et al. 2001).

All three of the ribosomal rRNAs are transcribed as a single transcript and processed as they are being transcribed. In the absence of RNases, there is an accumulation of cistronic transcript containing all three rRNAs (Apirion and Gegenheimer 1981). This allows for equal amounts of each rRNA to be synthesized. Since the rRNAs are co-expressed as a cistronic unit, their synthesis can be co-regulated, particularly during transcription initiation and elongation by NusA and NusB (S. C. Li, Squires, and Squires 1984). NusA and NusB interact with regions upstream of the first rRNA gene containing anti-terminator loops (box A-C) and hairpin sequences that control the rate of RNA polymerase speed (S. C. Li, Squires, and Squires 1984) (Figure 1).

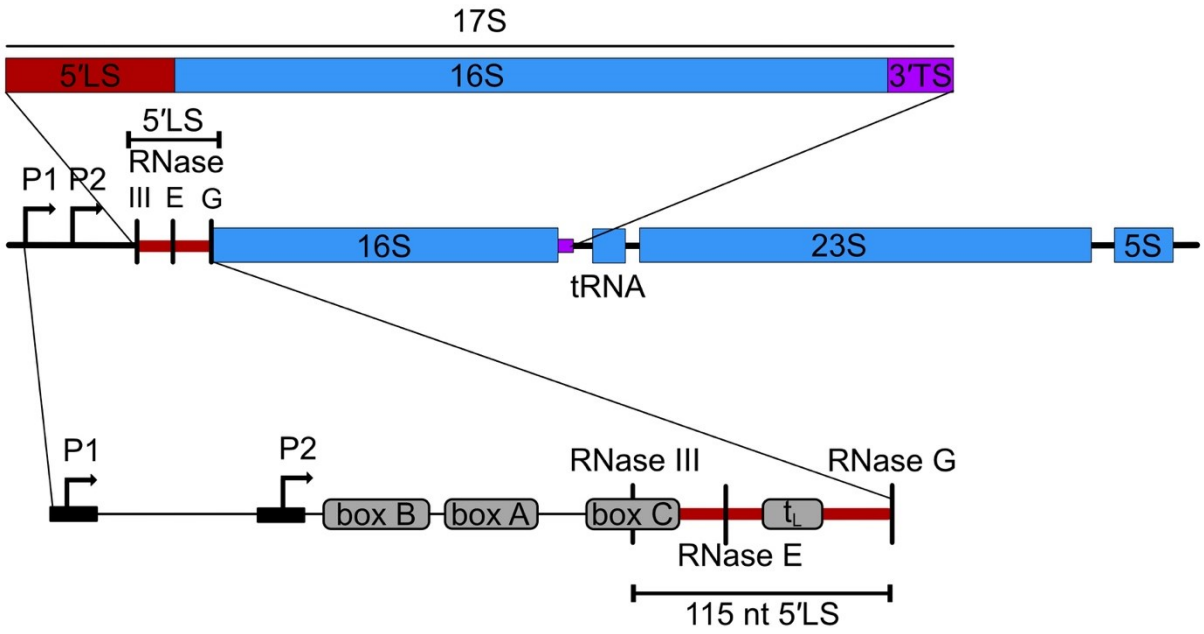


Figure 1: *E. coli rrn* operon

Structure of a typical *rrn* operon in *E. coli*. Blue boxes represent the coding regions of the 16S, tRNA, 23S and 5S. The cut sites for RNase III, RNase E, and RNase G are highlighted with vertical lines. The 115 nt 5′ leader sequence used in this study (5′LS) is highlighted by a red line and is located before the 16S sequence begins and starts at the RNase III cut site. The leader sequence containing boxes A, B, and C is removed by RNase III (S. C. Li, Squires, and Squires 1984). After the 16S coding sequence, a 33 nt 3′ trailer sequence in purple remains after RNase III cleavage. The combined 5′ LS, 16S, and 3′TS comprise the 17S rRNA. Two promoters, P1 and P2, respond to different signals within the cell (Sarmientos et al. 1983).

The *rrn* operons contain two promotor regions, P1 and P2, which have distinct mechanisms for responding to the metabolic state of the cell (Figure 1) (Sarmientos et al. 1983). Both P1 and P2 promoters are sensitive to the amino acid concentration within the cell, but the P1 promoter has a stronger response to the stringent response and the alarmone (p)ppGp (Sarmientos et al. 1983). By starving the cell of certain metabolites, ribosome synthesis can be halted at transcription initiation, and then restarted upon feeding to follow the kinetics of ribosome biogenesis *in vivo* (Hulscher et al. 2016). *E. coli rrn* operons are highly regulated to make sure that the proper amount of rRNA is available to match the growing needs of the cell, and to control the consumption of cellular energy and metabolites required to support ribosome biogenesis.

30S subunit Biogenesis and Maturation

During transcription, the nascent RNA begins to fold and is processed by ribonucleases (Figure 1). The rRNA genes are separated by intergenic regions that are recognized by ribonucleases and processing occurs in several steps to generate mature rRNA species. Once the full-length sequence of the 16S rRNA has been transcribed with its terminal sequence, a hairpin is formed between the leader sequence and the terminal sequence which is recognized and cleaved by RNase III (Nikolaev, Silengo, and Schlessinger 1973; Young and Steitz 1978). The resulting cleavage product is called the 17S rRNA which contains a 115 nt 5′ leader sequence (5′LS) and a 33 nt 3’ trailing sequence (3′TS). Following RNase III cleavage, the 17S rRNA is recognized by RNase E, which cleaves the 17S rRNA within the 5′LS generating a shorter 5′LS that contains 66 nt upstream of the 16S. Finally, RNase G cleaves at the very start of the 16S rRNA to generate the mature 5′ end (Z. Li, Pandit, and Deutscher 1999). The 3′TS was found to

be processed independently of the 5′LS by multiple RNases, including PNPase and YbeY (Sulthana and Deutscher 2013; Jacob et al. 2013).

The 16S rRNA is divided into four domains, the 5′ domain (nt: 1-561), central domain (562-916), 3′ major domain (917-1394), and 3′ minor domain (1395-1542.) As these domains are synthesized, they begin to fold with the help of r-proteins (Figure 2). Each domain is able to nucleate and fold independently of each other (Weitzmann et al. 1993). The RNA-RNA and RNA-protein contacts that occur during 16S folding, change as the molecules continues to be processed and matured (Stern et al. 1989). The first interaction a protein makes with the rRNA does not last, instead both change the others conformation as the 30S subunit matures. Some contacts that form initially refold as the 30S subunit takes form. An example of this is the folding of helix 3 in the presence of r-protein S4, where you can see the movement of helix 3 closer and farther to the protein, until it properly settles into its bound state (Kim et al. 2014).

Although the upstream leader sequence of the precursor 16S rRNA gets processed away, it is believed that parts of that sequence participates in co-transcriptional folding of the 16S rRNA by behaving in a chaperone like fashion (Besançon et al. 1999). Mutations that have been made to the upstream region affect the melting of the 16S rRNA, while showing the proper processing of the rRNA. The upstream RNA makes intermittent contacts with the 16S rRNA, sequestering specific sequences from participating in folding thereby separating structure formation spatially and temporally. Mutations in box C result in cold sensitivity and poor growth, suggesting that there are interactions between the upstream sequence and the 16S rRNA (Figure 1) (Schäferkordt et al. 2001). These upstream mutations result in the proper sequence of the RNA being transcribed, but it is misfolded. Ribosomes which contain misfolded rRNA are error prone when decoding mRNA and do not recognize stop codons, resulting in the buildup of

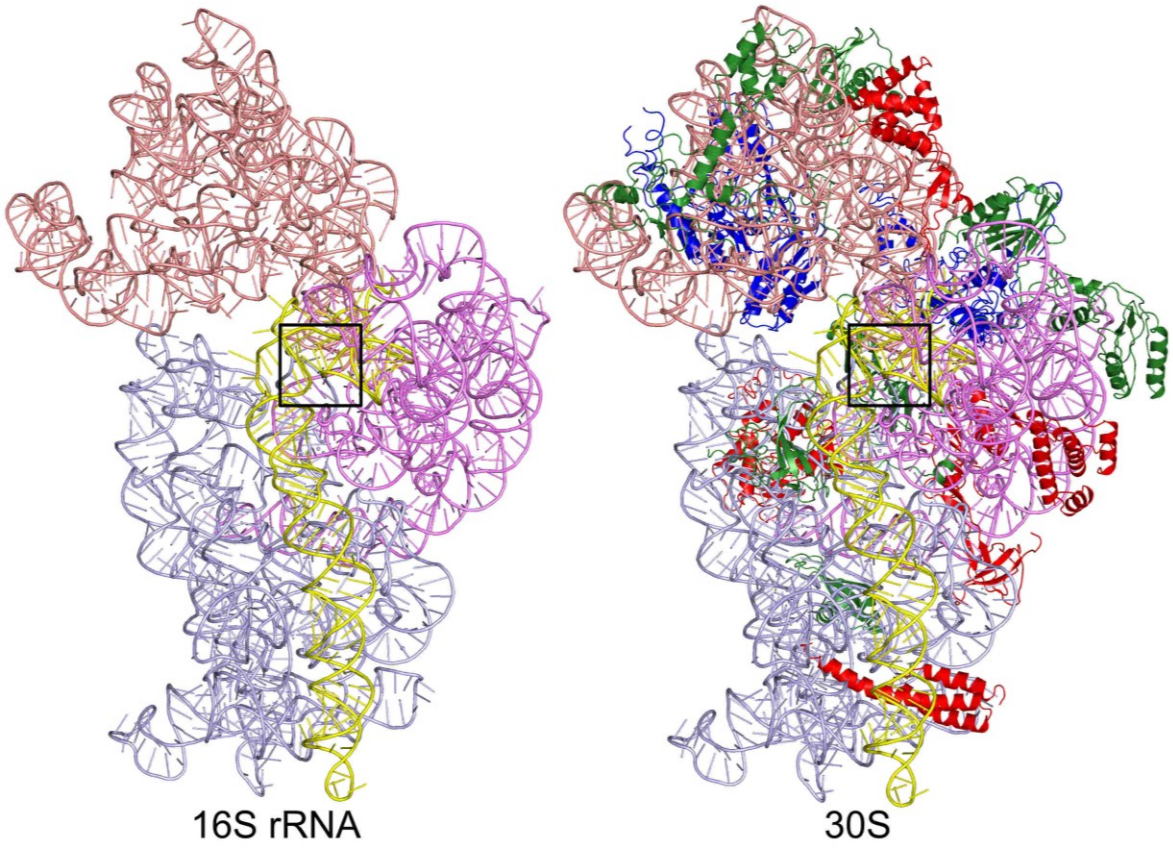


Figure 2: Tertiary fold of the 16S rRNA with the mature 30S subunit

On the left is the tertiary fold of the 16S rRNA without r-proteins shown (PDB: 4V7S) (Dunkle et al. 2010). The rRNA is color coded with the 5′ domain in blue, the central domain in purple, 3′ major domain in light pink and the 3′ minor domain in yellow. On the right is the mature 30S, with all of the r-proteins bound (PDB: 4V7S) (Dunkle et al. 2010). Primary proteins are colored in red, secondary proteins in green and tertiary proteins in blue. The position of the central pseudoknot, or helix 2, is highlighted by a square. It is positioned at the heart of the 30S, within the active site of the subunit.

toxic proteins (Roy-Chaudhuri, Kirthi, and Culver 2010). These examples demonstrate the co-transcriptional folding and guidance of RNA folding by proteins and RNA sequences not present in the mature 30S subunit.

Ribosomal proteins

Parts of the 16S rRNA are able to fold themselves *in vitro* in the presence of only salts and high temperature, but r-proteins are needed to stabilize the folded 16S rRNA in the cell (Adilakshmi, Ramaswamy, and Woodson 2005; Razi and Ortega 2017). The 21 different r-proteins (S1-21) have been categorized into three classes: primary, secondary and tertiary, based on their binding ability and pelleting with the 16S rRNA during *in vitro* 30S subunit reconstitution (Figure 2) (Mizushima and Nomura 1970; Held et al. 1974). The primary binding proteins bind directly to the rRNA, while secondary proteins can only bind when one or more primary proteins are already present. Tertiary binding proteins were classified as those that bind after a required heat step during *in vitro* reconstitution (Masayasu Nomura and Held 1974). The binding of each protein changes the fold of the rRNA, allowing for the next class of proteins to recognize, bind and cooperatively fold the rRNA (Talkington, Siuzdak, and Williamson 2005; Williamson 2006). Although the proteins bind in a hierarchical fashion, assembly and folding can proceed through a number of parallel pathways to generate a mature 30S subunit (Talkington, Siuzdak, and Williamson 2005; Adilakshmi, Bellur, and Woodson 2008; Gupta and Culver 2014).

Proteins are able to reorganize the rRNA as they bind and stabilize particular folds, while destabilizing non-native folds (Ramaswamy and Woodson 2009; Abeysirigunawardena and Woodson 2015). While S4 initiates nucleation by binding directly to the rRNA, S16 binds further away and destabilizes possible structures to allow for further on pathway folding of the 30S

subunit. Many ribosomal proteins have globular domains and disordered C-terminal tails. These tail penetrate the 16S rRNA and make contacts with them after they initially bind (Brodersen et al. 2002). As the pre-30S is matured, r-proteins that initially bind, change their structure, while changing reshaping the rRNA and vice-versa.

30S subunit reconstitution *in vitro*

Our earliest understanding of ribosome assembly came from efforts to reconstitute 30S subunits from their rRNA and protein components (M. Nomura 1968; Powers, Daubresse, and Noller 1993; Krzyzosiak et al. 1987). After purifying 30S subunits, it is possible to separate its components into native 16S rRNA and Total Proteins of 30S (TP30) which contains all the r-proteins of 30S subunit (Masayasu Nomura and Held 1974; Dodd, Kolb, and Nomura 1991). Then by combining and incubating the components together at 42 ℃, functional 30S subunits can be formed (Nierhaus 1990; Powers, Daubresse, and Noller 1993). Reconstitution can also be carried out using recombinantly purified variants of each r-protein (Culver and Noller 1999). However, it was found that if you combine all of the recombinant proteins together, the reconstitution is not as efficient as with native proteins in TP30 and results in nonfunctional molecules that do not migrate at the same size as mature 30S subunit (Culver and Noller 1999). The same effect can be seen if a ratio of 1:4 of 16S rRNA to TP30 or higher is combined, because the r-proteins bind the rRNA nonspecifically stabilizing nonfunctional folds. Instead of combining all the proteins together, recombinant proteins can be split into their classes (primary, secondary, tertiary) and incubated with the RNA sequentially. This allows for the hierarchical binding and cooperative folding of rRNA. Magnesium concentration and temperature (42 ℃) are required conditions for assembly of the tertiary proteins such that assembly can proceed from a ribosomal intermediate (RI) to RI* (Masayasu Nomura and Held 1974). RI particles sediment at

the same time as 21S particles, and the higher temperature allows for RNA rearrangements to occur resulting in RI* particles capable of binding tertiary proteins. Interestingly, at low temperatures, the RI accumulates *in vitro* indicating that there is a thermodynamic energy barrier that needs to be overcome by either proteins *in vivo* or high temperature *in vitro*.

Assembly Factors

In addition to 30S subunit ribosomal proteins, other proteins bind to the 30S ribosome during maturation but do not remain stably bound. These proteins are called assembly factors, which act to stabilize a particular fold or lower the activation energy for a particular structure and then must dissociate from the subunit before maturation is complete (Bunner et al. 2010; Woodson 2011). Assembly factors play an important role in promoting rapid and robust assembly of the ribosome under non ideal conditions. Assembly factors are not required for 30S subunit assembly *in vitro*, however they have been shown to aid assembly *in vitro* (Tamaru et al. 2018). Many assembly factors are non-essential for cell viability, but genomic deletion results in an increase in 17S rRNA, 30S and 50S subunits, decrease in 70S ribosomes, and cold sensitivity (Shajani, Sykes, and Williamson 2011). All of these characteristics are thought to arise from misfolded variants of 30S subunit, suggesting that proper folding is essential for not only ribosomal function but also for maturation. There are more than 17 documented assembly factors involved in bacterial 30S subunit ribosome assembly (Shajani, Sykes, and Williamson 2011).

RbfA and RsgA are two examples of assembly factors which play an important role in the maturation of 30S subunit. RsgA is a GTPase, capable of removing RbfA from maturing 30S subunit (Goto et al. 2011). Deletion of RsgA results in slow growth under non ideal conditions, due to RbfA remaining bound to 30S subunit and preventing the small subunit from forming functional 70S ribosomes. It also plays a checkpoint role in 30S subunit maturation, where it can

bind and flip out the A1493 nucleotide to check for its ability to discern codon-anticodon pairing (Razi, Guarné, and Ortega 2017). RsgA has also been found to bind and destabilize certain folds within the 30S subunit (both rRNA and r-proteins), allowing it to overcome kinetic barriers (Pedro Lopez-Alonso et al. 2017). It's possible for an assembly factor to use GTP to refold rRNA during maturation.

RbfA (ribosome-binding factor A) is capable of binding to immature and mature 30S subunits and has been suggested to play a role in the transition that occurs between the pre-30S into 30S subunit (Jones and Inouye 1996; Inoue et al. 2006). It was originally identified as an unknown protein in the operon of initiation factor 2, but later established as an assembly factor (Sands et al. 1988). Because of its ability to rescue a cold sensitive mutation at the 5′ end of the 16S rRNA, which was proposed to interfere with processing of the 5'LS, RbfA was predicted to bind directly to the mature 30S subunit (Cruz 1995). RbfA binds close of the 5′ end of 16S rRNA and is hypothesized to stabilize helix 1 using its C-terminus tail, allowing RNase G to cleave the 5′LS (Datta et al. 2007). RbfA binds to only 30S subunit and is not found in 70S or polysome fractions, hinting at it being a late stage assembly factor. Further work has highlighted RbfA's role in late stage 30S subunit maturation since 30S subunits in strains lacking RbfA contain regions of misfolded RNA indicated by exposed nucleotides (Soper et al. 2013). Once RbfA completed its goal, RsgA binds and removes RbfA. IF3, an initiation factor, is also able to remove or outcompete RbfA from mature ribosomes, highlighting RbfA's role as a barrier between immature and mature 30S subunit and between 30S subunit biogenesis and translation initiation (I.M. Sharma and S.A. Woodson, personal communication).

Chapter 2: Twister Ribozyme

Introduction

Ribozymes are catalytically active ribonucleic acids whose functions arise from their properly folded conformations (reviewed in (Jimenez, Polanco, and Lupták 2015)). A number of classes of small self-cleaving ribozymes have been discovered (Jimenez, Polanco, and Lupták 2015), with a recent addition of the twister ribozyme family (Roth et al. 2014). The twister class was identified bioinformatically and its biological function is yet to be uncovered (Roth et al. 2014). The twister ribozyme family is highly conserved and is dispersed throughout the tree of life. It is classified by a secondary fold containing three to four helices, two internal loops and one outer loop (Figure 3) (Roth et al. 2014; Liu et al. 2014). Its tertiary fold is created by two pseudoknots, T1 and T2 (Figure 3, labelled in orange and blue). These pseudoknots play a role in positioning nucleotide G39 to act as a base within the active site, allowing for an in-line attack by the 2′ O of U-1 on the adjacent phosphodiester, while the A1 base acts to stabilize and protonate the leaving product (Wilson et al. 2016) (Figure 3B). The active site is conformationally flexible, reaching the active state after local conformational changes that properly set up the inline attack (Gaines, Giese, and York 2019). Because all the crystal structures involve a non-reactive nucleotide in the active site, the splayed conformation of the phosphate backbone and the in-line attack position of the 2′OH is not observed, the proper conformational state of the active site has been difficult to deduce. Once its tertiary structure has formed, twister self-cleaves between the U-1 and A1 nucleotide (nt), resulting in a short 5′ cleavage product (6 nt) and a 3′ product containing the majority of the ribozyme (48 nt). The short length of twister ribozymes has made them a

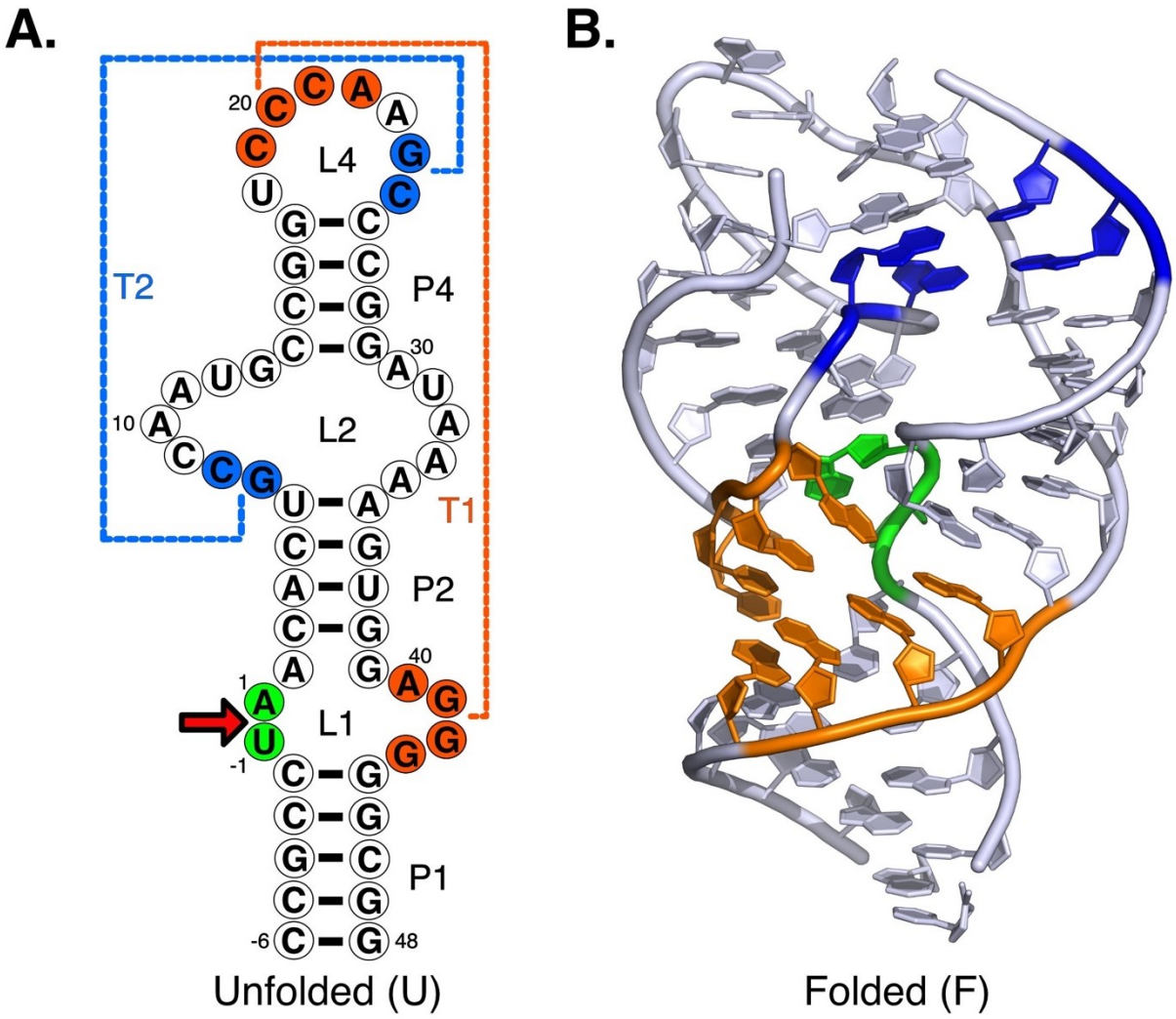


Figure 3: Secondary and tertiary structure of twister ribozyme

A) The secondary structure of twister ribozyme from *Oryza sativa* (54 nt), labeled using the universal numbering system (Roth et al. 2014; Wilson et al. 2016). The three helices that are common within the twister family are labelled P1, P2, and P4. The two internal loops are labelled L1 and L2, with the third outer loop labelled L4. The two pseudoknots are color coded in orange and blue, and labelled T1 and T2 respectively. In its proper tertiary fold, the nucleotides in the active site are positioned for cleavage between the U and A (bright green), as indicated with a red arrow in A.

B) The crystal structure of *Oryza sativa* twister ribozyme (PDB: 4OJI) (Liu et al. 2014), color coded as in A.

great model system for studying RNA folding, as they can be easily manipulated while maintaining high activity *in vitro*.

Previous work has focused on the activity of twister ribozymes by studying how cleavage and folding rates are dependent on pH and divalent metal ion concentrations (Wilson et al. 2016; Gebetsberger and Micura 2017). Because of its ability to detect the dynamics of RNA, single-molecule Fluorescence Resonance Energy Transfer (smFRET) spectroscopy is a direct way to study ribozyme folding and cleavage. This has been previously accomplished by labeling twister ribozymes with fluorophores and then following the dynamics of twister ribozyme as it folds and unfolds, without undergoing cleavage due to the substitution of a deoxy ribose (Panja et al. 2017b; Vušurović et al. 2017). These studies on non-cleavable forms of the RNA highlight the structural rearrangements that occur during folding, such as the formation of specific pseudoknots and the stabilization of helix P1 (Vušurović et al. 2017).

Because RNA begins to fold as it is being synthesized, Hua et al. mimicked co-transcriptional folding of twister ribozyme with a helicase (Hua et al. 2018). The helicase was used as a way to over-come the difficulties of using an RNA polymerase for transcription of the RNA and mimicked the synthesis of twister ribozyme from the 5′ to 3′ end. This work showed the early structural rearrangements that occur while the twister ribozyme tries to reach its energy minima. The probability of misfolding was higher during its maiden folding transition, while during equilibrium the probability of misfolding decreases (Hua et al. 2018). This suggested a "molecular memory" for the twister ribozyme: the probability of misfolding is high for the first folding event, due to the mispairing of the secondary structure. However, once the proper secondary structure has been achieved, the tertiary structure has a lower probability of misfolding.

All of the above-mentioned studies suffered from a common challenge that arises from studying self-cleaving ribozymes, which is that self-cleavage cannot be controlled. In order to prevent the ribozyme from cleaving itself during folding experiments, a 2′ deoxyribose was substituted at U-1, rendering the RNA non-reactive. This allows for folding and unfolding, but the dynamics of the uncleavable twister may not be representative of the structural rearrangements that occur immediately prior to cleavage. It is difficult to predict how sufficient these folded states are for the activity of the ribozyme because it is incapable of forming product.

One way to investigate the dynamics and the resulting product of twister ribozyme folding is to use a photo-active chemical group to separate secondary from tertiary fold controlling the cleavage of twister ribozyme (reviewed in (Klán et al. 2013)). Photocaged compounds have been used previously to study both protein and RNA interactions, by preventing them from forming the required interactions (Buck et al. 2007; Panja et al. 2015). Once activated with a microsecond exposure to light, isomerization of the protecting group results in loss of the arylcarbonylmethyl group and regeneration of the proper RNA base.

In collaboration with the Greenberg lab, who synthesized a modified photocage which is activated by visible light, I prepared a caged *Oryza sativa* twister ribozyme and used this RNA to examine tertiary folding of the cleavable ribozyme in real time. I found that immobilized photocaged RNA is activated by a 405 nm laser within 30 milliseconds with spatial precision. I demonstrated that it is possible to use a photochemical group in combination with smFRET spectroscopy to study the folding and cleavage rate of twister ribozyme. By having both spatial and temporal control of RNA dynamics, we can study the connection between folding and function in a more relevant biological and kinetic context.

Results

Characterizing the cleavage activity of a fluorescent twister ribozyme

Previous studies were done using a commercially available internal Cy3 residue, which destabilized native structure, raising its midpoint of folding when compared to unlabeled twister ribozyme (Panja et al. 2017a; Hua et al. 2018). In collaboration with the Greenberg lab, twister ribozyme with Cy3 attached to 2′ OH of U18 was synthesized, because this attachment of the dye was expected to be less disruptive to the RNA structure, as seen by other labs (Vušurović et al. 2017). To measure the self-cleavage activity of the Cy3-labeled twister ribozyme, the RNA was incubated at different $MgCl_2$ concentrations (0.02-5 mM) for 1 min at 20 ℃. The RNA products were resolved on a 10% denaturing gel. Figure 4A and 4B show that increasing $MgCl_2$ concentration results in more product after 1 min of reaction, as expected. Because twister ribozyme is activated by Mg^{2+} ions during enzymatic ligation of synthetic RNA fragments, a significant fraction of cleaved product (~ 30%) is present in the control lane (Figure 4 A,B). When the remaining fraction cleaved versus [Mg^{2+}] was fit to a cooperative folding model, the midpoint for populating the active conformation of the ribozyme was 0.19 (±0.065) mM $MgCl_2$, three times higher than the 0.065 mM $MgCl_2$ midpoint for self-cleavage of unlabeled twister ribozyme (Panja et al. 2017a).

At 2 mM $MgCl_2$, the fraction of twister ribozyme cleaved within 1 min reached saturation (~ 75% cleavage), making it a good condition to study its rate of cleavage. To measure the rate of cleavage, twister ribozyme was incubated at 2 mM $MgCl_2$ for a range of times (Figure 4C). After 70 seconds, the maximum fraction cleaved (70%) was reached. The fraction of product was then fit using a single exponential equation to give an observed rate constant of k_{obs} = 0.06 (±0.004) s^{-1} or 3.6 min^{-1} (Figure 4D). This self-cleavage rate was similar to what others have

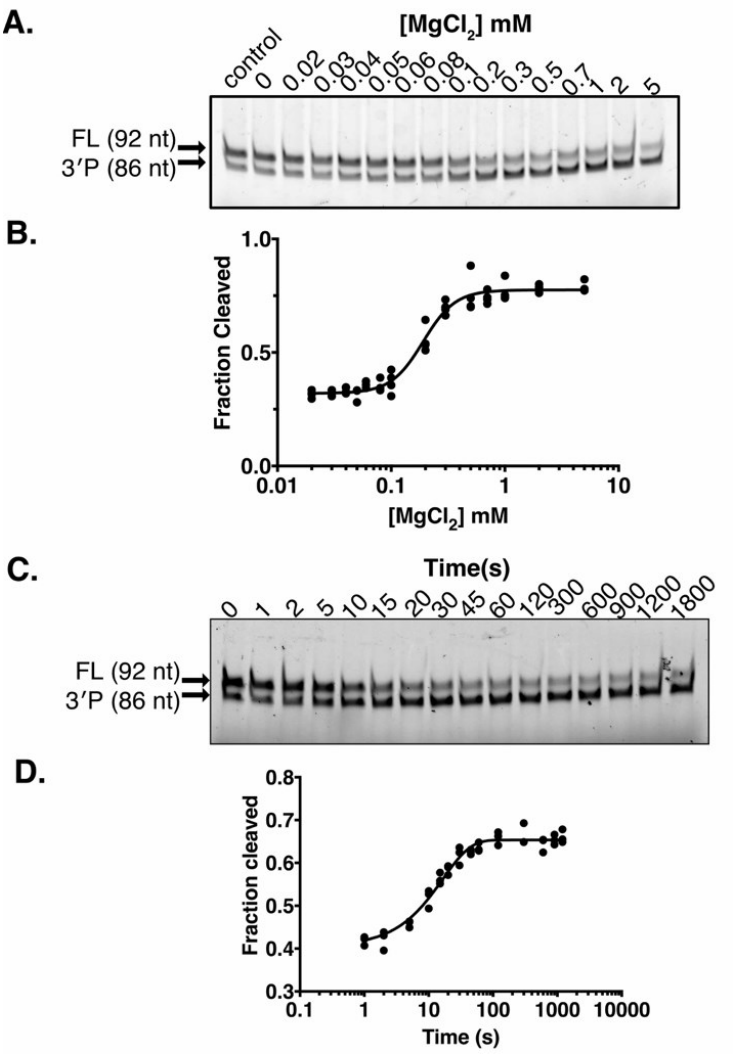


Figure 4: Self-cleavage of U18-Cy3 labelled twister ribozyme.

A) A denaturing 10% polyacrylamide gel showing self-cleavage of the extended (SA5) twister ribozyme after 1 min in the presence of a range of [$MgCl_2$]. The RNA is labelled with Cy3 at U18 2' OH.

B) Fraction of cleaved twister ribozyme as in (A). Results of four repeats were globally fit to a Hill equation, with a midpoint ($[Mg^{2+}]_{1/2}$) of 0.19 ± 0.019 mM $MgCl_2$.

C) A denaturing 10% gel showing the self-cleavage at different times after adding 2 mM $MgCl_2$. Cleavage results in an 86 nucleotide 3′ product.

D) Fraction of cleaved product versus time after 2 mM $MgCl_2$ was added to start the self-cleavage reaction. All four repeats are globally fit to a single exponential rate equation, yielding an observed rate constant k_{obs} = 0.06 ± 0.004 s^{-1}, or 3.6 min^{-1}.

reported; *env22* twister had a rate of k_{obs} = 2.44 min^{-1}, and *O. sativa* k_{obs} = 2.45 min^{-1} at pH 7.0 and k_{obs} = 3.5 min^{-1} at pH 8.5 at 10 mM $MgCl_2$ and room temperature (Wilson et al. 2016; Ren et al. 2014; Gebetsberger and Micura 2017).

The midpoint of cleavage of the 2′ OH labelled twister ribozyme in this study has a lower midpoint of folding (0.19 mM $MgCl_2$) than the previously studied internally labeled twister ribozyme (1.4 mM $MgCl_2$), but still somewhat perturbing when compared to non-fluorescently labelled RNA (0.065 mM $MgCl_2$) (Panja et al. 2017a). This difference is due to the position of the Cy3 fluorophore, since the twister ribozyme used previously had the Cy3 fluorophore inserted into the backbone of the RNA between U18 and G17, which led to the destabilization of the T2 pseudoknot. Crystal structures show that U18 and A23 interact, resulting in G24 and C25 bulging out, priming them to form the T2 pseudoknot. The backbone addition of Cy3 interrupts the proper positioning for the U18-A23 base pair formation. I hypothesize that changing the Cy3 position to the 2′ OH group on the ribose of U18 alleviated the strain induced by the backbone-linked fluorophore and is therefore a more biologically relevant ribozyme construct to study coupling of folding and cleavage activity.

Detecting the folding and unfolding of twister ribozyme via smFRET

Although gel-based assays measure the $MgCl_2$ –dependence and rate of cleavage, the dynamics of folding are hidden due to cleavage being an indirect measurement of folding. Single-molecule fluorescence resonance energy transfer (smFRET) spectroscopy allows us to directly look at the conformational dynamics of RNA molecules and draw conclusions about the kinetics of twister ribozyme folding and cleavage. To adapt the twister ribozyme for single-molecule studies, as previously done by our lab, the 3′ end was extended with an SA5 sequence which anneals to a DNA containing a biotin at the 5′ end allowing for surface immobilization

(Figure 5A) (Panja et al. 2017a; Hua et al. 2018). The DNA oligonucleotide is also labeled with a Cy5 acceptor fluorophore at the 3′ end in order to observe FRET from the Cy3 donor fluorophore located on U18 of the ribozyme. Once the extended twister RNA was annealed to the Cy5-labeled, biotinylated DNA oligonucleotide, it was then immobilized on a quartz microscope slide, using neutravidin, and passivated with DDS (Hua et al. 2014). Molecules were then exposed to either a green (532 nm) or red (633 nm) laser to excite Cy3 (the donor) or Cy5 (acceptor), respectively. As the molecule folds and unfolds, the distance between the Cy3 and Cy5 fluorophores changes, which results in a change in resonance energy transfer. I measure this as an anti-correlated change in the fluorescence intensity of Cy5 and Cy3. Detecting molecules fold and unfold in the presence of different $MgCl_2$ concentrations will allow me to see how $MgCl_2$ changes the folding dynamics and potentially detect new conformations via FRET values that arise from different structural changes while twister ribozyme folds.

In order to look at the folding and unfolding rates under different $MgCl_2$ concentrations, a non-cleavable variant of the twister ribozyme was made, containing a deoxyuridine (dU) at U-1. This chemical substitution removes the 2′OH of the uridine (rU) which is the nucleophile for the self-cleavage reaction. This substitution was used previously to generate crystal structures, and to detect the folding and unfolding rates (Figure 5A) (Liu et al. 2014). Each movie recorded the dynamics of the dU twister ribozyme over five minutes. The movies produced trajectories of the intensity of the Cy3 and Cy5 fluorophores for single molecules in the field of view. The change in FRET efficiency, E_{FRET}, was calculated from these intensities after subtracting the background and correcting for cross-talk (Figure 5B).

From the trajectories, we can see there are two states that are populated, a low FRET state (FRET = ~ 0.1) and a high FRET state (FRET = ~0.9) (Figure 5B). In the low FRET state, the

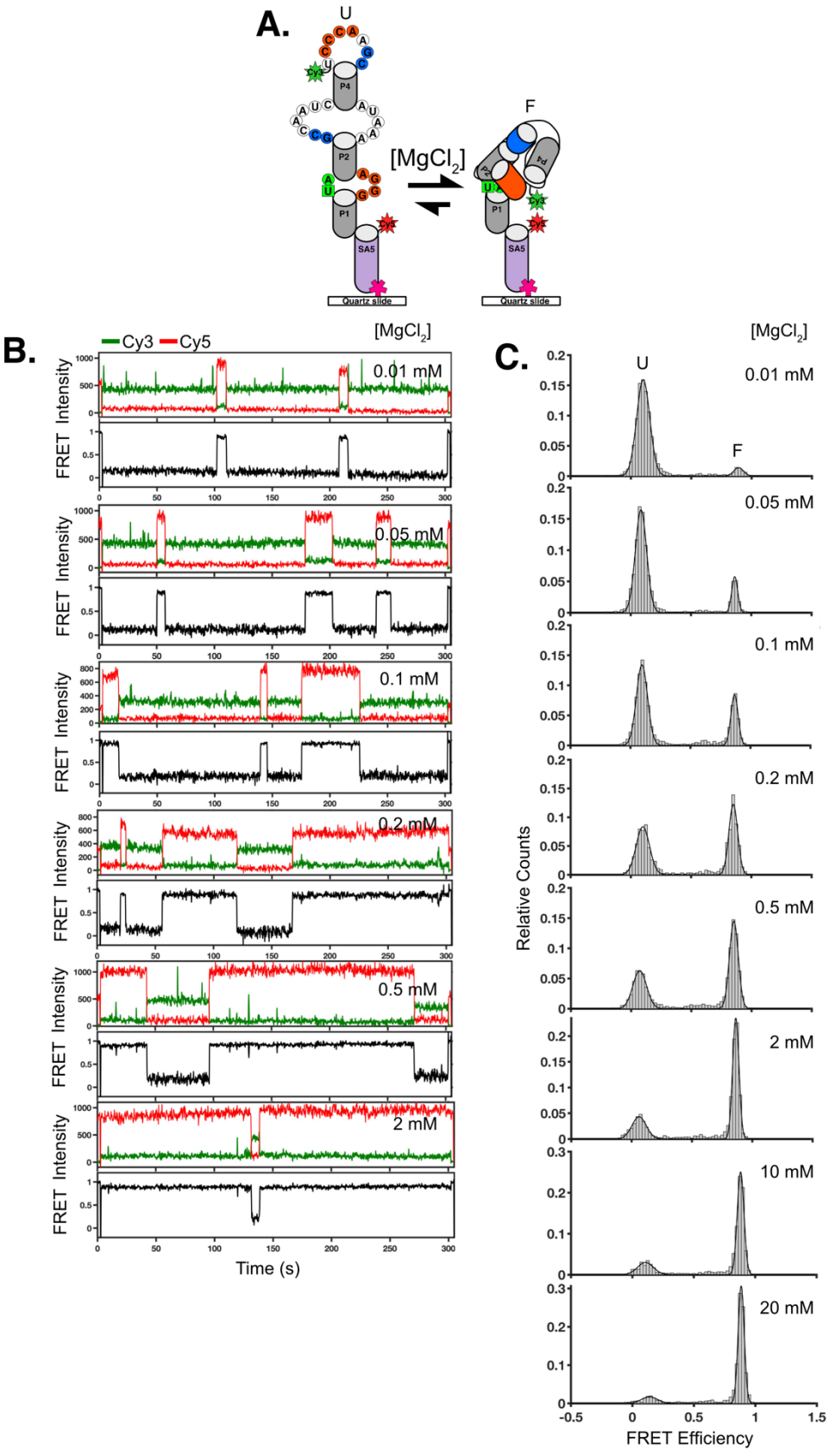

A.
U
F
[MgCl2]
P4
P2
P1
SA5
Quartz slide
B.
Cy3
Cy5
[MgCl2]
0.01 mM
0.05 mM
0.1 mM
0.2 mM
0.5 mM
2 mM
FRET Intensity
Time (s)
C.
[MgCl2]
U
F
0.01 mM
0.05 mM
0.1 mM
0.2 mM
0.5 mM
2 mM
10 mM
20 mM
Relative Counts
FRET Efficiency

Figure 5: Folding dynamics of non-cleavable twister ribozyme in a range of $MgCl_2$ concentrations

A) Diagram of single molecule folding experiment using non-cleavable (dU –1) twister ribozyme. The extended twister ribozyme with Cy3 attached at U18 2′OH was annealed to a DNA oligonucleotide containing Cy5 at the 3′ end and biotin at the 5′ end, forming the SA5 helix (violet). At different $MgCl_2$ concentrations, twister ribozyme remains folded (high FRET state) or unfolded (low FRET state) for different lengths of time.
B) Examples of 5-minute traces recorded for different $MgCl_2$ concentrations from 0.01 mM to 2 mM. Cy3 and Cy5 signals are represented by green and red respectively, while the calculated FRET value is represented by black traces. As $MgCl_2$ concentration increases, the time spent by molecules in the high FRET increases.
C) Population histograms showing the change in the folded and unfolded

Cy3 and Cy5 dyes are far apart, as has been seen previously for the unfolded secondary structure of twister ribozyme (Hua et al. 2018; Panja et al. 2017a). The high FRET state is the result of twister ribozyme folding into its tertiary fold, as Cy3 and Cy5 come close together upon pseudoknot formation (Figure 5A). As the $MgCl_2$ concentration increases, the dwell time that twister ribozyme spends in the folded state increases from ~ 10 s at 0.01 mM $MgCl_2$, to greater than 5 min (the duration of the movie) above 2 mM $MgCl_2$, while the lifetime of the unfolded state decreases (Figure 5B). This is to be expected, as Mg^{2+} stabilizes the compact tertiary fold of RNAs, by allowing their negatively charged backbones to come closer together without repelling each other (reviewed in (Draper, Grilley, and Soto 2005)). By stabilizing these long-range interactions, Mg^{2+} increases the length of time that the molecule spends in the folded state, increasing the probability of local rearrangements resulting in a product in ribozymes.

FRET data from thousands of molecules was used to build population FRET histograms containing the states of all the molecules (Figure 5C). The population histograms show a much clearer picture of the effect of $MgCl_2$ on the state of twister ribozyme. As $MgCl_2$ is increasing, the high FRET, folded population increases while the low FRET, unfolded population decreases. There are also no detectable intermediate conformations, indicated by the lack of density in the histogram between the low and high FRET peaks. This suggests that folding of twister ribozyme can be approximated as a simple two state system, with an unfolded and a folded state, as previously reported (Panja et al. 2017a). The lack of intermediates can either be because our time resolution of 50 ms is not fast enough to capture the formation of intermediate folds, or because these intermediates do not form under these conditions.

Rate of folding and unfolding of dU twister ribozyme

Population histograms are a great way to detect intermediates and understand the folding pathway of RNA. However, in order to get the rates of folding and unfolding between states, the time duration of the unfolded and the folded state needs to be measured. Time trajectories can be used to calculate rate constants for folding and unfolding by measuring the dwell times of the unfolded state, representing the folding rate, and the dwell time of the folded state, representing the unfolding rate. Instead of performing this operation by hand for hundreds of molecules at different conditions, Hidden Markov Modelling fits the FRET trace for each molecule to a system with two or more states to generate an idealized trace with transition times between low and high FRET states, which are then used to calculate the dwell times.

I used a two-state HMM to analyze the folding kinetics of the dU twister ribozyme (Figure 6A). Increasing the amount of $MgCl_2$ slightly increases the rate of folding (k_F), while dramatically decreases the rate of unfolding (k_U) (Figure 6B). At 0.02 mM $MgCl_2$, the rate of unfolding is ~10 times faster than the rate of folding (k_U is 0.16 s^{-1} and k_F is 0.014 s^{-1}). Therefore, the unfolded state dominates the kinetics under this condition. If a molecule does fold, it will quickly unfold and remain unfolded because the probability of folding is very low. However, at 0.1 mM $MgCl_2$, the rates intersect at 0.022 s^{-1}. This midpoint can also be visualized in the population histograms as an equal amount of area under the curve of each peak (Figure 5C). As $MgCl_2$ concentration increases, the unfolding rate continues to drop and the folding rate dominates, as the molecules continue to remain folded. After 0.2 mM $MgCl_2$, it becomes difficult to measure rates of folding or unfolding because the chance of a molecule switching between states within the duration of the experiment decreases, as the lifetime of the folded state approaches the timeframe of the experiment. Also, because HMM analysis requires transition

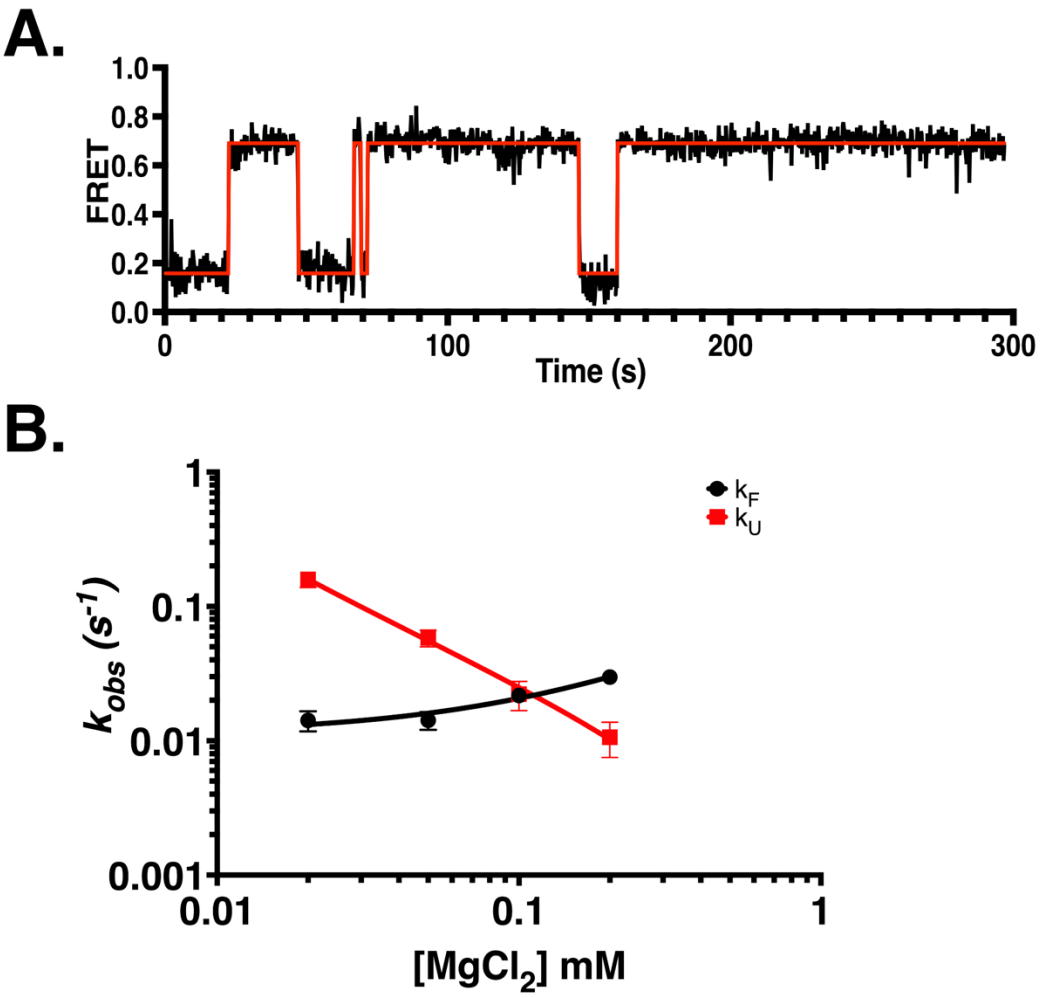


Figure 6: Dependence of folding and unfolding rates on $MgCl_2$ concentration for the non-cleavable twister ribozyme.

A) An example of a FRET trace in black, analyzed using hidden Markov modeling in red.
B) Rate constants for folding (k_F, black) and unfolding (k_U, red) of twister ribozyme in different $MgCl_2$ concentrations. Rate constants for folding and unfolding were both fit to a modified Hill equation, $k_{obs}=[Mg^{2+}]/(k + [Mg^{2+}])+c$, where k is the rate constant. As $MgCl_2$ concentration increases, the rate of unfolding decreases and the rate of folding increases.

events, the molecule selection can skew towards the more active molecules and not the stably folded or unfolded molecules within the time frame of the movie. This can explain why the rates intersect at 0.1 mM $MgCl_2$, and not 0.16 mM $MgCl_2$, as the slower transitioning molecules are not counted.

Previous data collected on the internally Cy3 labelled dU twister ribozyme also showed that the unfolding rate decreases with $MgCl_2$, while the folding rate increases with increasing $MgCl_2$ (Panja et al. 2017a). At the lowest $MgCl_2$ (0.5 mM), $k_U = 2.0\ s^{-1}$ and a $k_F = 0.01\ s^{-1}$; these rate constants become equal at 0.1 mM $MgCl_2$ with $k_U = k_F = 0.07\ s^{-1}$. The rate of unfolding decreases dramatically, after being ~200 times larger than the folding rate at 0.5 mM $MgCl_2$, while the folding rate increases relatively slower by 7 times. From both of the data sets, irrespective of the labeling of the twister ribozyme, increasing $MgCl_2$ destabilizes the unfolded state rapidly, and makes the folded state favorable. Similar correlation can be seen in the hairpin ribozyme, where with increasing $MgCl_2$, the active docked state of the ribozyme is more favorable, and the undocked state less favorable (Tan et al. 2003). The rate of switching also decreases with $MgCl_2$, as the hairpin ribozyme remains docked at high $MgCl_2$ concentration. Although the $MgCl_2$ does not play a role in the reaction, it does stabilize the active conformation, just as seen in the twister ribozyme.

Effect of $MnCl_2$ and PEG8000 on twister ribozyme folding

Twister ribozyme is able to fold and self-cleave in the presence of different divalent metals. Mg^{2+} is the second most prominent ion in the cell, behind K^+, and is known to support RNA and rRNA folding in the cell (Draper, Grilley, and Soto 2005). However, it was found that twister ribozyme has a lower midpoint of cleavage in the presence of transition metal ions such as Mn^{2+} and Ni^{2+}, which has the lowest midpoint of catalysis (Panja et al. 2017a). In order to test

the effect of $MnCl_2$ on folding, I titrated the dU twister ribozyme with $MnCl_2$ (Figure 7A). Just as with $MgCl_2$, the folded population increases with $MnCl_2$, however, the midpoint of folding is 0.064 ± 0.0040 mM, two and a half times less than the folding midpoint in $MgCl_2$ (0.16 ± 0.021 mM) (Figure 7C). This decrease in the midpoint of folding highlights the ability of Mn^{2+} to stabilize the folded state at a lower concentration (Figure 7C). The Hill constant is also larger for $MnCl_2$ (n = 1.6). This larger cooperativity of folding also indicated that the folded state is more stable in $MnCl_2$, compared to $MgCl_2$ (n = 1.1) which folds noncooperatively.

Divalent metal ions are not the only cellular components which support the folding of RNA. Molecular crowding agents, such as PEG8000, are known to favor the folded, or condensed, structures in both RNA and proteins (Kilburn et al. 2010; Dupuis, Holmstrom, and Nesbitt 2014). With the cell itself being a crowded environment, molecular crowders are helpful in simulating cellular conditions in which an RNA would natively fold. To see the effect of crowders on twister ribozyme folding, the twister RNA was titrated with $MgCl_2$ in the presence of 10% PEG8000 (Figure 7B). At 0.02 mM $MgCl_2$ in the presence of PEG8000, there is a higher population of folded molecules (28% folded) compared to $MgCl_2$ only (6.2% folded) (Figure 7B). The midpoint of folding in the presence of 10% PEG8000 drops to 0.087 ± 0.029 mM $MgCl_2$ (Figure 7C). The most prominent effect is that PEG8000 supports more molecules folded under lower $MgCl_2$ concentrations, but has little effect at higher concentrations, as the folded population is similar with (75% folded) and without (74% folded) the crowder present (Figure 7C). This could mean that the rate of folding is influenced the most by crowders when there is less Mg^{2+} present, as the excluded volume effect is supporting more of the folded state (Kilburn et al. 2010). However, when $MgCl_2$ concentration increases, the effect of the divalent ion overshadows the crowders contribution to folding. Due to this, the Hill constant drops to 0.67, as

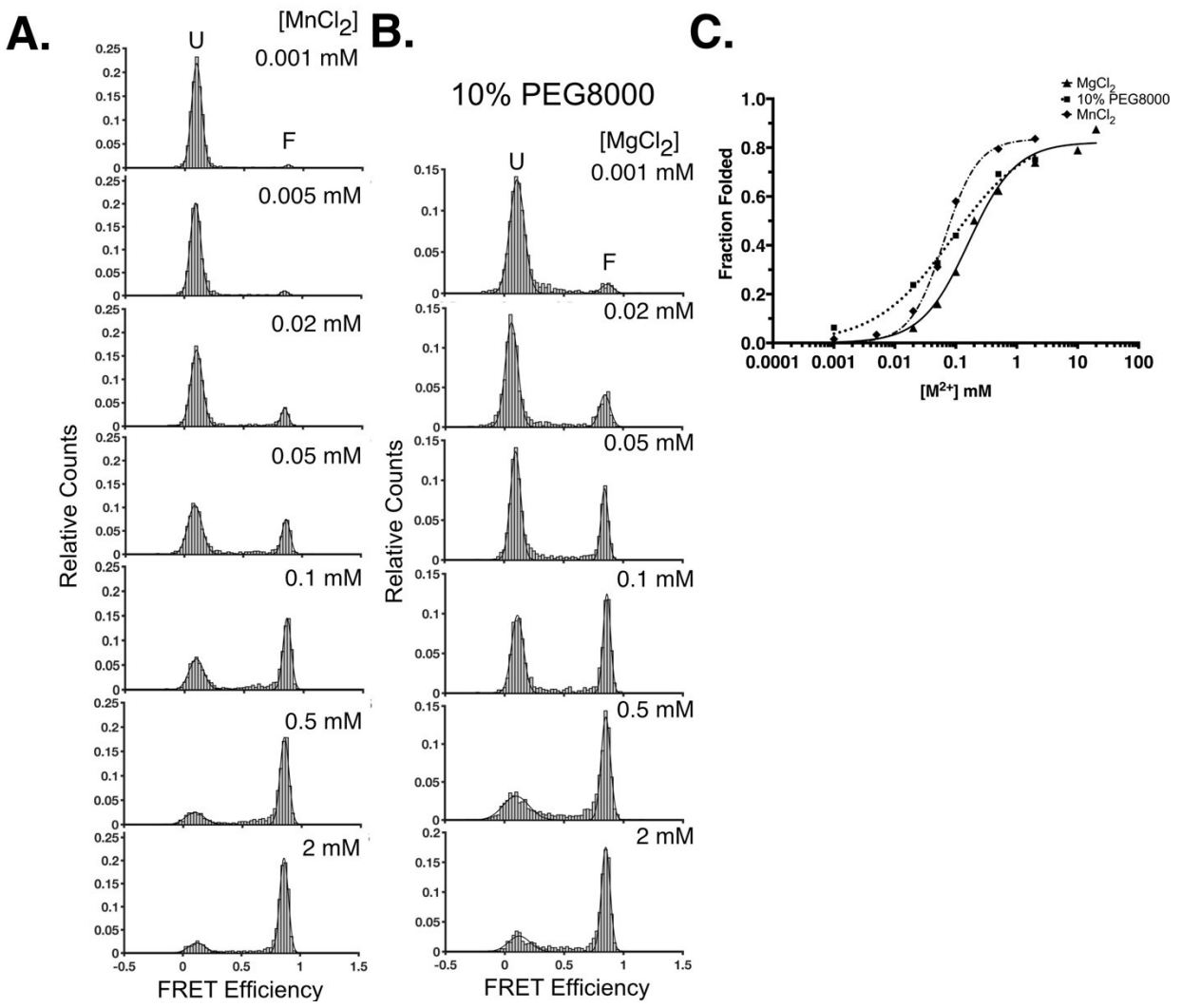


Figure 7: The effects of $MnCl_2$ and PEG8000 on twister ribozyme folding

A) Population histograms of non-cleavable (dU) twister ribozyme under a range of $MnCl_2$ concentrations. As the concentration of $MnCl_2$ increases, the amount of folded (high FRET) twister ribozyme increases more steeply than in $MgCl_2$. Each histogram is fit to a double Gaussian function.

B) Population histograms of non-cleavable (dU) twister ribozyme under a range of $MgCl_2$ concentrations in the presence of 10% PEG8000. A larger fraction of twister molecules is folded in lower $MgCl_2$ concentrations in the presence of PEG than in the absence of PEG. Each histogram is fit to a double Gaussian function.

C) Fraction of folded non-cleavable twister ribozyme as a function of $[Mg^{2+}]$, $[Mg^{2+}]$ with 10% PEG8000, and $[Mn^{2+}]$. All three curves are fit to the Hill equation. The midpoints of folding of twister ribozyme are 0.16 ± 0.021 mM $MgCl_2$ with $n = 1.1 \pm 0.15$, 0.087 ± 0.029 mM $MgCl_2$ in 10% PEG8000 with $n = 0.69 \pm 0.12$, and 0.064 ± 0.0040 mM $MnCl_2$ with $n = 1.6 \pm 0.15$.

$MgCl_2$ is contributing less to the stability of the tertiary structure due to the effect of PEG8000 at lower $MgCl_2$ concentrations.

My data support a model in which different metal ions and crowders support the tertiary fold of twister to a higher extent than $MgCl_2$ alone. Crowders can affect folding by collapsing the molecules together, which would increase the probability of the pseudoknots forming. However, it's still unknown why Mn^{2+} and other metal ions are preferred over Mg^{2+}. Neither component seems to be stabilizing any distinct intermediate conformation with an intermediate FRET that can be detected with my choice of fluorophores (Figure 7A,B). What I can deduce from these experiments is that $MnCl_2$ influences both the folding and unfolding rate at low and high concentration, while PEG8000 has the most influence at lower concentration of metal ions where it destabilizes the unfolded state (Kilburn et al. 2010).

rU twister ribozyme remains folded once it is cleaved

Testing the non-cleavable twister ribozyme is important to deduce the rates of folding and unfolding under different concentrations of $MgCl_2$. However, I wanted to look at the dynamics of rU twister ribozyme which can self-cleave. To characterize the folding dynamics after cleavage, the wild type rU ribozyme was first pre-incubated in 20 mM $MgCl_2$ to make sure all of the twister ribozyme had been cleaved. In the presence of 2 mM $MgCl_2$, after being cleaved in 20 mM $MgCl_2$, the majority of the population was in the folded state, with very few molecules remaining in the unfolded state (Figure 8). In order to test the stability of this fold, the cleaved twister ribozyme was challenged with 5 mM EDTA to sequester any $MgCl_2$ in the reaction that might support the tertiary interactions. To my surprise, the majority of the population stayed in the folded, high FRET state as indicated in Figure 8. Even in 50 mM EDTA, the ribozyme remained folded (Figure 8). In fact, 50 mM EDTA had destabilized the small

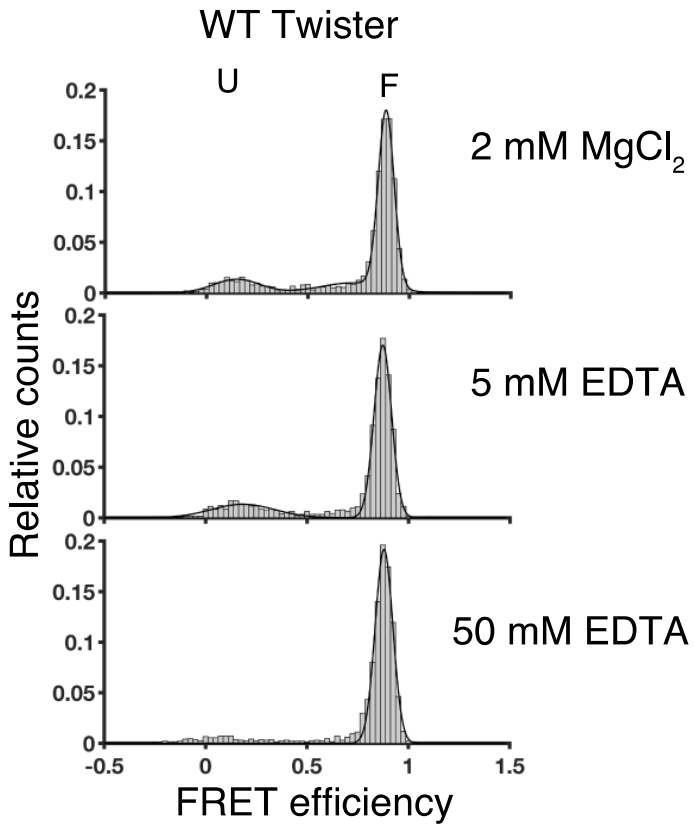


Figure 8: Twister ribozyme remains folded once it is cleaved.

Wild type (cleavable) twister ribozyme was allowed to self-cleave, and then imaged in the presence of 2 mM $MgCl_2$, 5 mM EDTA, and 50 mM EDTA. The majority of the population was in the folded state, irrespective of the effective $MgCl_2$ concentration.

population of unfolded or intermediate species that were present, suggesting perhaps that free $MgCl_2$ was stabilizing the low FRET state observed in 2 mM $MgCl_2$ and 5 mM EDTA (Figure 8). The increased EDTA also did not change the FRET value observed of the folded state. The FRET value remained close to 0.9, showing that the folded structure remained the same. If there was a change, it was not being observed due to the positioning of the fluorophores. The effect of EDTA on the folded structure also highlights that the twister ribozyme undergoes a structural rearrangement which permits it to remain in the high FRET conformation. It is possible that the was a change, it was not being observed due to the positioning of the fluorophores. The effect of EDTA on the folded structure also highlights that the twister ribozyme undergoes a structural rearrangement which permits it to remain in the high FRET conformation. It is possible that the monovalent potassium present in the buffer (100 mM) also supports the folded state in the presence of EDTA. Stabilization of the folded state of the cleaved twister ribozyme had been reported previously for the *env22* twister variant (Vušurović et al. 2017).

Generating a photoactivatable twister ribozyme

Studying the dynamics of wildtype twister is difficult due to its fast folding and self-cleavage. Previously, controlling the tertiary fold was done by flowing in $MgCl_2$ during the experiment, with folding driven by the presence of divalent ions (Panja et al. 2017b; Hua et al. 2018; Vušurović et al. 2017). However, this method is difficult, inconsistent, time consuming, and requires a large amount of RNA. Flow in experiments also require 1-2 seconds of deadtime for equilibrations, as new components are being flown in. In order to overcome these obstacles, the twister ribozyme was made into a photoactivatable RNA, which can start its cleavage process in the presence of light. The photocage on the RNA confines the molecule in the unfolded state by blocking the formation of pseudoknot T2, until it is activated by light.

Previously published data have shown that a mutation of G24 destabilizes pseudoknot T2, resulting in a majority of molecules remaining in the unfolded state at high $MgCl_2$ (Panja et al. 2017a) (Figure 3). Mutation of G42 within pseudoknot T1 has a similar effect, leaving most of twister ribozymes unfolded (Figure 3A). A protecting group that blocks base pairing at these positions would allow the photocage to behave as a "conditional mutation", by disrupting one of the pseudoknots from forming Watson-Crick base pairs until deprotected by light.

In collaboration with Huabing Sun, I synthesized the photo-caged, extended twister ribozyme by ligating synthetic oligonucleotides with the desired chemical modifications (Figure 9). Due to the inefficiency of coupling a photocaged ribonucleotide phosphoramidite in the middle of a synthesized RNA, it was decided to divide the twister sequence into two fragments, with the photocage attached to G24 near the 3′ end of the second fragment. Because the RNA with photocaged G24 did not react with light as expected (see below), we prepared a second RNA in which the photocage was kept at pseudoknot T2 but was attached to G7, located on the opposite side of the pseudoknot (Figure 9). This modification blocks formation of the T2 pseudoknot and therefore prevents the cleavage of twister ribozyme.

Activation of the photocage attached to twister ribozyme in bulk

This newly synthesized photocaged RNA is based on previously established p-hydroxylphenacyl group (Klán et al. 2013), with a modification which allows it to be activated by blue light with a wavelength of 405 nm (Figure 10A). The chemically modified guanosine with the protecting group at O6 was chemically synthesized by Huabing. Although the guanosine monomer was deprotected by 405 nm light, it was still unknown if it would be deprotected when incorporated into a larger RNA molecule. Twister ribozyme with the photocage attached at G7 (G7PC), was exposed to 405 nm light in solution (Figure 10B) in a light box containing blue

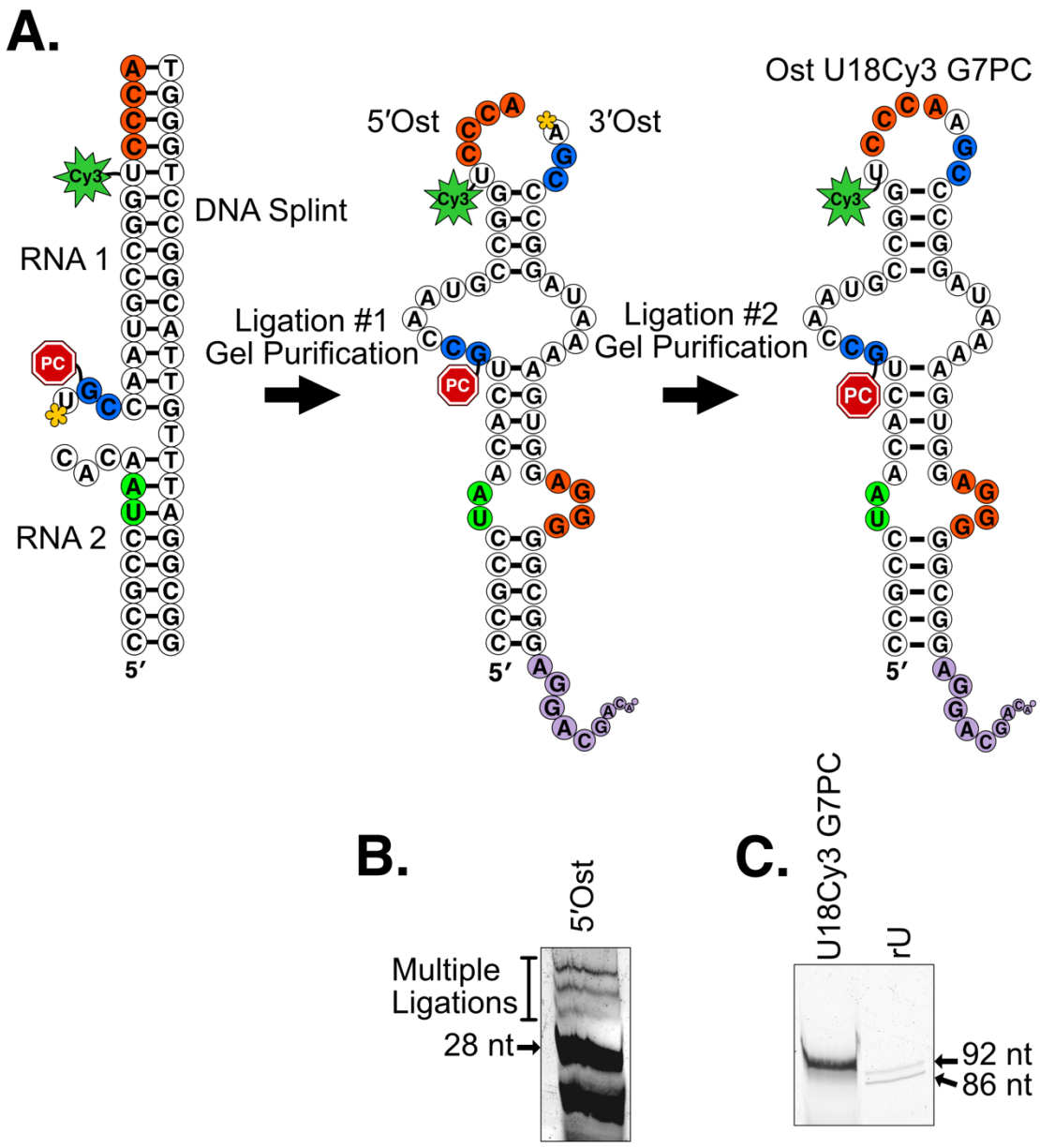


Figure 9: Preparation of twister ribozyme U18Cy3 G7PC by ligation

A) Schematic of the ligation process of twister U18Cy3 G7PC**.** RNA1 was synthesized in house with the photocage and labelled with Cy3 dye. It was ligated to phosphorylated (yellow star) RNA2 to make the 28 nt long 5′Ost RNA. Then it was annealed to the 64 nt 3′ Ost RNA-DNA hybrid containing the rest of the twister ribozyme sequence (RNA) and the SA5 sequence (DNA).

B) Gel purification of 28 nt 5′Ost, product of the first ligation. Bands resolved on a 16% denaturing gel. Above 28 nt are multiple ligation products. Band below is a contaminant.

C) The ligated product was resolved on a 10% denaturing gel, next to WT twister RNA that is cleaved. Final photocaged product is 92 nt in length, and no product visible.

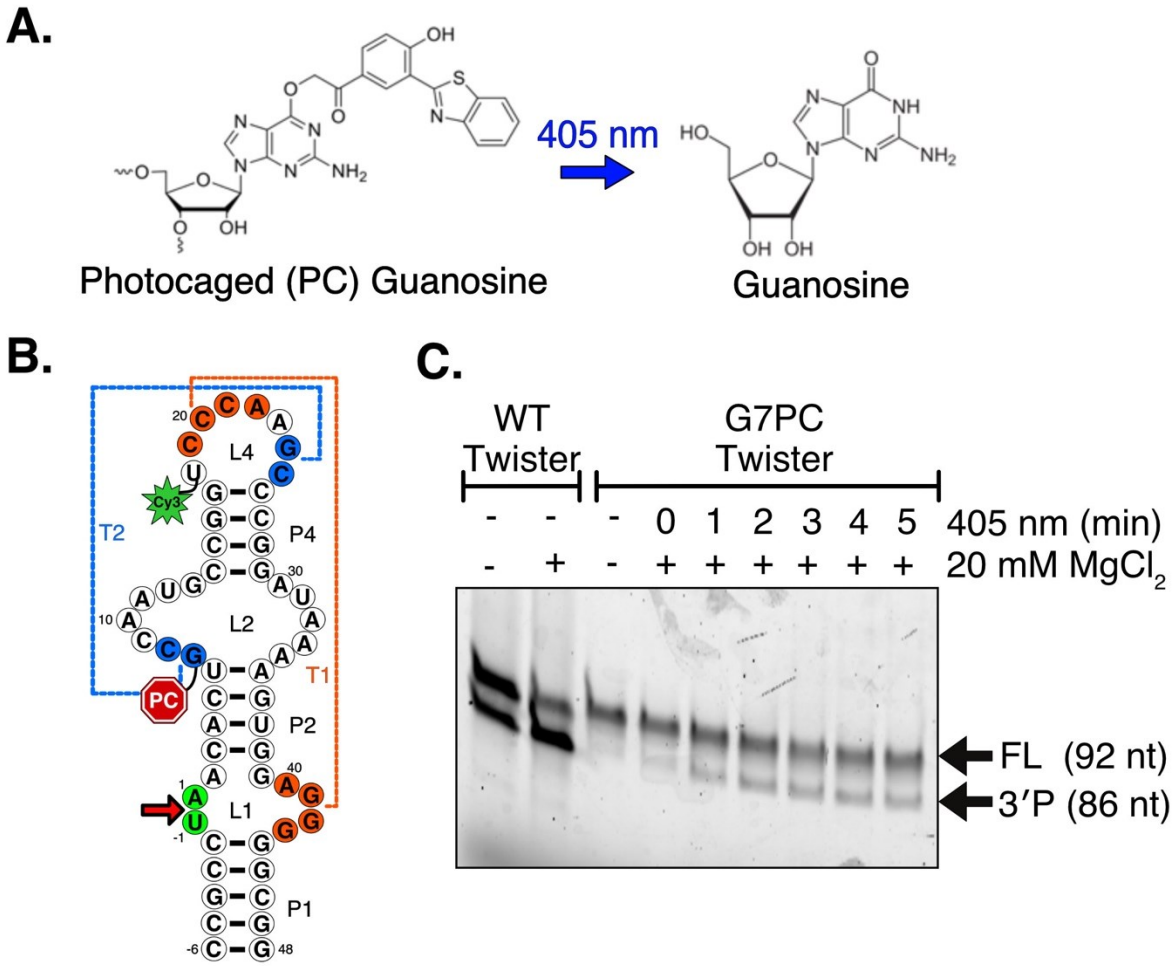


Figure 10: Photocaged guanosine allows rapid ribozyme activation with 405 nm light

A) Chemical structure of photocaged (PC) guanosine being activated by 405 nm light (blue arrow) and the final keto form of non-photocaged guanosine.

B) Secondary structure of *Oryza sativa* (rice), with stem P1, P2, and P4, and two internal loops, L1, L2 and an outer loop, L4. Cy3 is attached to the 2′ OH of U18. The photocage protecting group (PC) is attached to the C6 carbonyl of G7. The orange and blue nucleotides form pseudoknots T1 and T2, which are necessary for self-cleavage between U-1 and A1 in green, as indicated by the red arrow.

C) Photocaged twister self-cleaves only after activation with 405 nm light. Twister RNA in bulk was exposed to 405 nm light for 0-5 minutes, then incubated with 20 mM $MgCl_2$ for 1 min and quenched with denaturing dye (7M Urea + 50 mM EDTA). The samples were resolved on a 10% denaturing gel and scanned for Cy3 intensity. Two species were detected, the full length (FL) twister-SA5 RNA, 92 nt long, and the self-cleaved 3′ product (3′P) 86 nt long. After 5 minutes of blue light exposure, 30% of the RNA was converted to 3′ product.

LEDs. Without being exposed to blue light, there was no cleavage product in the control without $MgCl_2$, and in the 0 min time point with 20 mM $MgCl_2$ (Figure 10C). However, with increased time of exposure to blue light, the product increased in the presence of $MgCl_2$. The amount of cleaved twister ribozyme increased from zero, to 30% after 5 min of exposure. This shows that the photocage prevents cleavage and can be activated by 405 nm light as part of a larger RNA, even with diffuse light which does not penetrate the interior of the tubes and liquid effectively.

Photocage can be effectively activated by 405 nm light in single molecule experiments

The benefits of using this newly designed photocage are its ability to be activated by light within the visible spectrum, microsecond response times, it is scarless when activated (it produces the proper keto tautomer of guanosine) and gives temporal control of RNA folding. Because the gel-based assay required a five-minute exposure to reach 30% cleavage, this created doubt about whether it would be quickly activated. However, the photocage was tested in smFRET experiments with the idea that the poor yield is due to the diffuse, non-specific activation that occurs within a light box. To confirm that the photocage did not interfere with folding after activation, G7PC-dU twister ribozyme was used to perform analogous smFRET experiments as with the non-photocaged dU twister ribozyme. Before activation with 405 nm light, most of the molecules remained in the unfolded state (Figure 11B). However, once the molecules were exposed to 405 nm light for 0.25 seconds, a large population of molecules attained the folded state as evidenced by the appearance of bright Cy5 spots in the FRET channel (Figure 11B). Notably, focused activation by the 405 nm laser confined the uncaging reaction to defined fields of view (FOV) thereby allowing for imaging of multiple fields of view in a single slide channel. This further confirms that reappearance of folding dynamics is dependent on

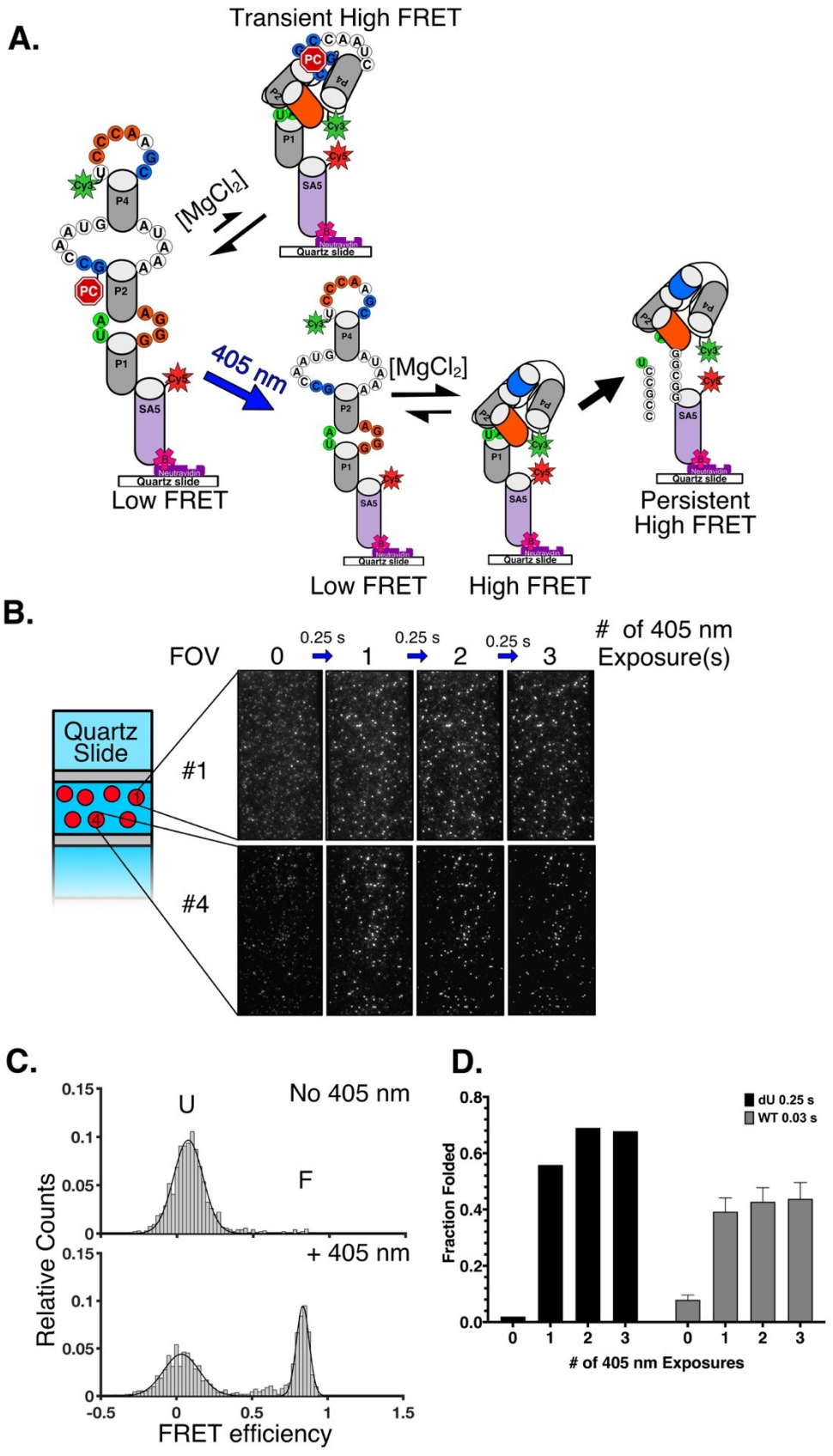

A.
Transient High FRET
[MgCl2]
405 nm
[MgCl2]
Low FRET
Low FRET
High FRET
Persistent
High FRET
Quartz slide
B.
FOV
0
0.25 s
1
0.25 s
2
0.25 s
3
# of 405 nm
Exposure(s)
Quartz
Slide
#1
#4
C.
No 405 nm
U
F
+ 405 nm
Relative Counts
FRET efficiency
D.
dU 0.25 s
WT 0.03 s
Fraction Folded
# of 405 nm Exposures

Figure 11: Photocage enables efficient RNA activation during smFRET experiments

A) Twister is annealed to an oligonucleotide (SA5) containing Cy5 at the 5′ and biotin at the 3′ end, which tethers twister to the quartz slide via neutravidin in purple. In the presence of $MgCl_2$ the photocaged G (PC, red octagon) obstructs the formation of pseudoknot T2 (blue), resulting in a stable low FRET signal. At high $MgCl_2$, twister ribozyme can transiently sample the folded state. After the photocage is activated by 405 nm light, twister ribozyme is able to test the folded state, and eventually self-cleave, forming a persistent high FRET product.

B) The photocage is activated only within the currently observed field of view (FOV), ~75 µm × 37 µm. Each FOV, diagramed by red circles on a single channel within the slide chamber, was exposed three times to 405 nm light (0.25 s each). Without any blue light exposure, there are very few bright spots in the Cy5 channel with the green laser on. After 0.25 sec of exposure to blue light, depicted with a blue arrow, molecules become bright as they fold into a high FRET state.

C) At 2 mM $MgCl_2$, photocaged twister RNA remains in its unfolded, low FRET state. After 0.25 sec of 405 nm light exposure, ~56% of molecules go into their high FRET, folded state.

D) Most molecules are uncaged after the first exposure to 405 nm light. Further blue light exposure does not increase the number of folded molecules. dU twister molecules were exposed for 250 ms to 405 nm and rU twister molecules were exposed for 30 ms.

activation of the photocage specifically by quick exposure to 405 nm light, resulting in a scarless RNA.

Next, I wanted to quantify the population of molecules which fold after activation by blue light. Molecules from multiple FOVs were recorded before 405 nm exposure, and then molecules were recorded after 405 nm exposure. Even in 2 mM $MgCl_2$ or higher, at the point which the non-photocaged dU twister ribozyme plateaued in the folded state, the photocaged molecules remained in the unfolded state, with virtually no tertiary folded molecules (Figure 11C). After 0.25 s of 405 nm exposure, a large population of folded molecules appeared. About 56% of the molecules folded, while in the non-photocaged dU twister ribozyme tested previously, 74% of the population was folded. This discrepancy in the amount of folded population may represent a small fraction of photocaged molecules that cannot be activated by light. To test the response of the photocage to increasing amounts of 405 nm light, multiple fields of views were exposed to blue light, three times each. The largest percentage of the photocage molecules are activated by the first exposure to 405 nm, with 56% of molecules folding. Upon another 0.25 s exposure of the same FOV, there was a modest increase in the folded population (up to 69%) which did not increase further with additional exposures (Figure 11D). A similar method was used with the self-cleaving rU twister ribozyme, but each exposure was for 0.03 s (Figure 11D). The rU twister ribozyme showed a suppressed response, with the major folding event occurring after the first 405 nm exposure, with 40% of the molecules folded, and remained steady with subsequent activations. This shows that the photocage can be activated by a single, rapid exposure of 0.03 s to light. The inability of a larger population of rU twister ribozyme to fold after activation is not due to the quick exposure to light, but instead is due to the structural variation in the RNA as will be discussed later.

Although not all the molecules are activated by 405 nm exposure, the population histograms highlight that the unfolded structure is equal to the structure observed in the non-photocaged dU twister ribozyme, and the folded state after photocage activation is also equal to the folded state of the dU twister and rU twister ribozymes. The photocage is able to "cage" the proper secondary fold of twister and allow for the proper tertiary fold after it is activated. In the presence of the photocage, the molecules are in the unfolded low FRET state (~0.1), then after activation, the molecules are able to organize into the tertiary folded, high FRET state (~0.9), and then go on to self-cleave while in the high FRET state (Figure 11A). Non-photocaged rU twister experiments have shown that once the RNA is cleaved, it is represented by a persistent high FRET state (Figure 8).

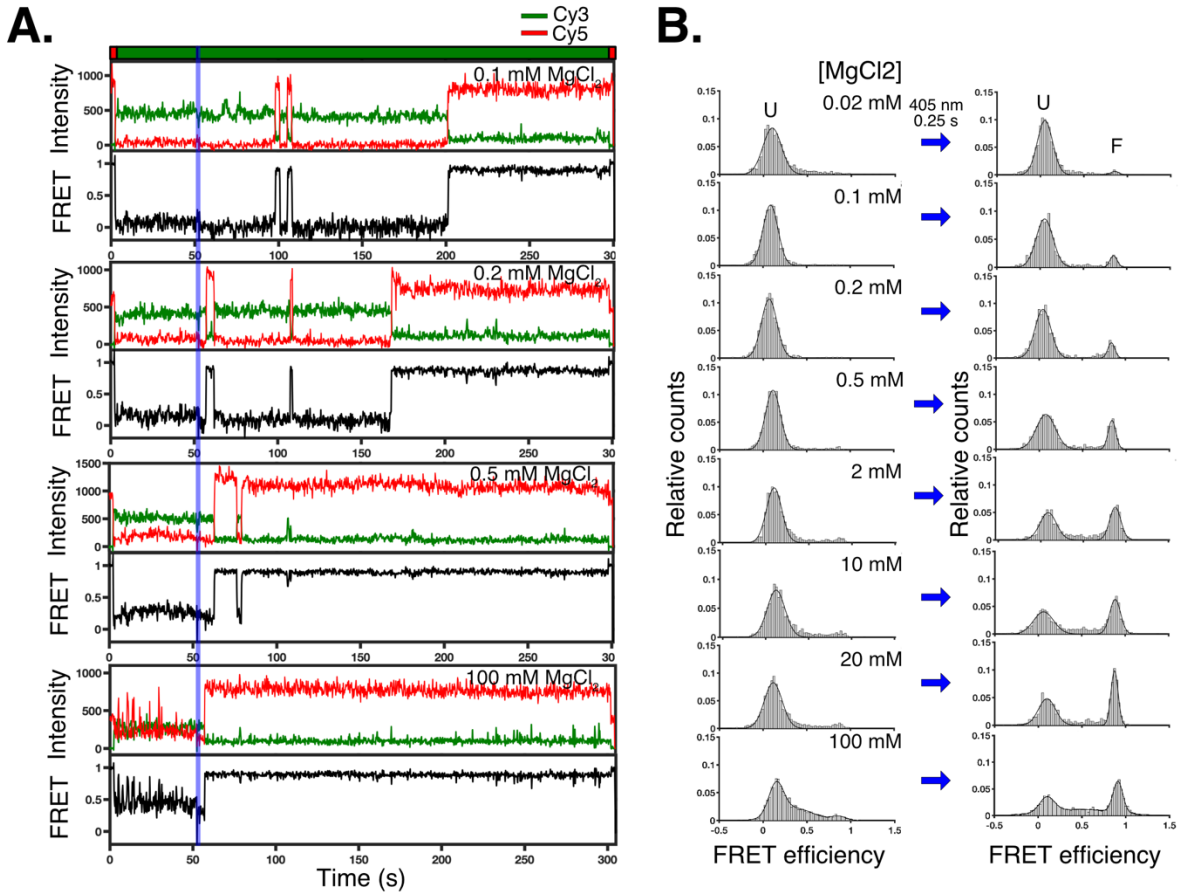
Detecting the folding dynamics of photocaged rU twister ribozyme

After verifying the activation of the photocage by blue light and folding of twister ribozyme into the proper tertiary fold, the dynamics of folding and cleavage of the photocaged rU twister ribozyme was analyzed. Five-minute movies were recorded at different $MgCl_2$ concentrations within the transition range fraction folded, and the ribozyme behavior before and after 405 nm exposure was compared. Within the initial 52.5 s of the movies for $MgCl_2$ concentrations up to 2 mM MgCl, all molecules remain in a static low FRET state indicating that they are unfolded due to the photocage (Figure 12A). In 100 mM $MgCl_2$, the molecules rapidly test the folded state in the presence of the photocage but cannot remain stably folded due to the presence of the photocage inhibiting formation of one of the pseudoknots.

Twister ribozyme samples the folded state before cleavage occurs at low Mg concentrations

Following activation, the photocaged rU twister ribozyme undergoes transitions from the low to high FRET states. These transitions often lead to a persistent high FRET state lasting ≥ 30

s which does not transition back to the low FRET state within the duration of the experiment (Figure 12A). I interpret the persistent high FRET to represent a cleavage event, since the cleaved twister ribozyme exhibits a long-lived static high FRET state (Figure 8). Other labs have also reported that the tertiary fold is stabilized after cleavage in twister ribozymes from a different species (Vušurović et al. 2017). When compared to the photocaged dU twister ribozyme, it becomes clearer that these persistent high FRET states are the result of cleavage (Figure 13A). The dU twister ribozyme does not remain in the folded state; it transitions between the folded and unfolded states, similar to what I saw in the non-photocaged dU molecule (Figure 5A).


A.
Cy3
Cy5
0.1 mM MgCl2
0.2 mM MgCl2
0.5 mM MgCl2
100 mM MgCl2
Intensity
FRET
Time (s)
B.
[MgCl2]
0.02 mM
0.1 mM
0.2 mM
0.5 mM
2 mM
10 mM
20 mM
100 mM
405 nm
0.25 s
U
F
Relative counts
FRET efficiency

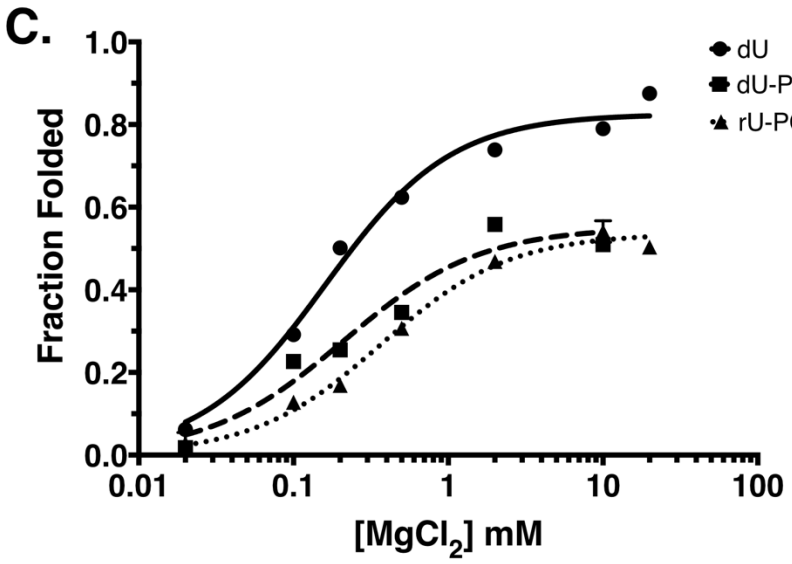

C.
dU
dU-PC
rU-PC
Fraction Folded
[MgCl2] mM

Figure 12: Light activation of twister ribozyme reveals the coupling between folding and self-cleavage

A) Sample trajectories of rU twister under different $MgCl_2$ concentrations. Red and green traces represent the Cy5 and Cy3 signal respectively, and the black trace represents the calculated FRET. Blue line represents time of 405 nm exposure for 0.25 s. Red and green boxes at the top represents the sequence of lasers, 10 frames of red at the start and end of the movie, and green laser in the middle.
B) Population histograms of rU twister before and after 405 nm light exposure in a range of $MgCl_2$ concentrations.
C) $MgCl_2$ dependent folding of the different twister variants tested. The high FRET population, f_N, was fit to a Hill equation, $f_F = [Mg^{2+}]n/([Mg^{2+}]_{1/2}{}^n + [Mg^{2+}]^n)$, in which $[Mg^{2+}]_{1/2}$ is the midpoint and n is the gradient, $\partial f_N/\partial[Mg^{2+}]$. dU twister without the photocage has a midpoint of 0.16 ± 0.024 mM $MgCl_2$; dU twister with the photocage (dU-PC) after activation has a midpoint of 0.21 ± 0.084 mM $MgCl_2$; rU twister with photocage (rU-PC) after activation has a midpoint of 0.37 ± 0.056 mM $MgCl_2$.

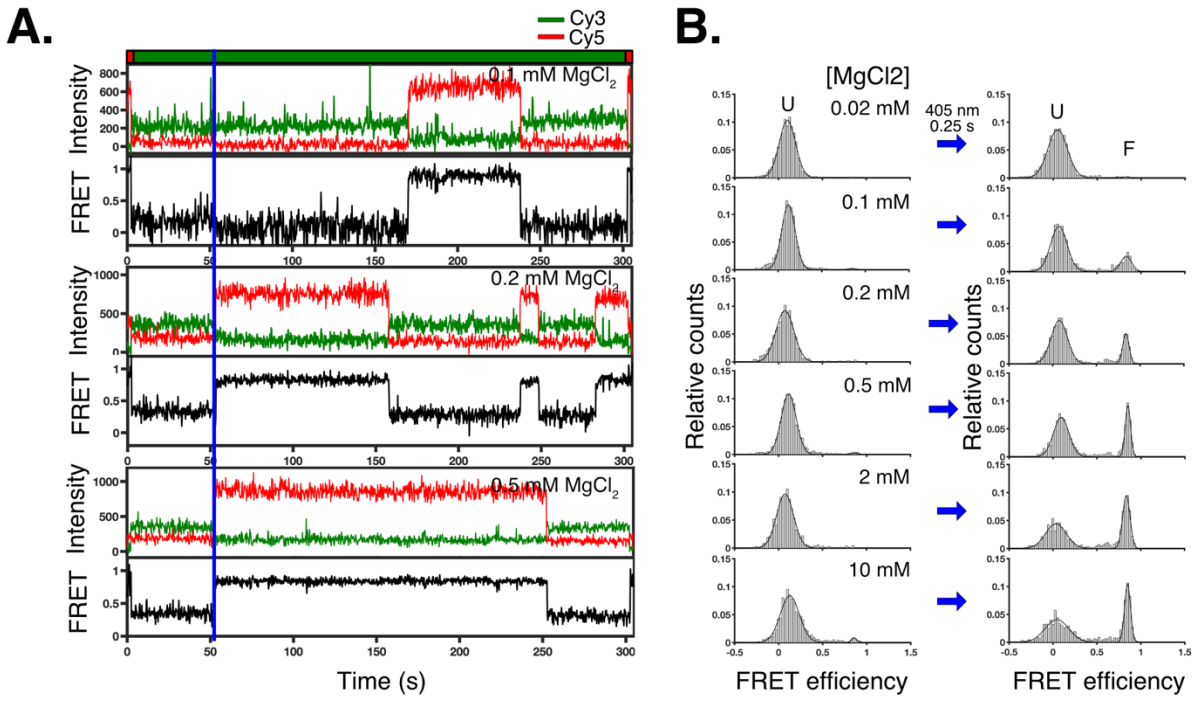


Figure 13: Activation of photocaged non-cleavable (dU) twister ribozyme.

A) Sample trajectories under different $MgCl_2$ concentrations of non-cleavable (dU) photocaged twister ribozyme. Just as in figure 12A, red and green traces represent the Cy5 and Cy3 signal intensity, respectively, and the black trace represents the calculated FRET efficiency.

B) Population histograms of non-cleavable twister ribozyme before and after 405 nm light exposure in a range of $MgCl_2$ concentrations.

At low $MgCl_2$ concentrations, twister ribozyme is more likely to fold and unfold multiple times, before attaining the persistent fold, highlighting how twister can transiently fold without self-cleaving as previously reported (Panja et al. 2017b). At 100 mM $MgCl_2$, twister ribozyme folds into its persistent state within a few seconds after the photocage is activated by 405 nm light, showing that RNA in higher concentration of $MgCl_2$ are less likely to unfold due to their low rate of unfolding, and go straight into tertiary fold and self-cleavage.

Interestingly, at 100 mM $MgCl_2$, there appears to be a population with intermediate FRET. This FRET species is present with and without the photocage, but slightly decreases after 405 nm light exposure. It is possible that these represent new intermediate structures which are only sampled in high concentrations of $MgCl_2$. However, upon inspection of individual trajectories, it appears that the lifetimes of the intermediate states are very short, often only lasting a single frame. Furthermore, molecules that sample the intermediate states after activation with the 405 nm laser never fold into the stable high FRET state. This behavior is more consistent with molecules which have not been uncaged.

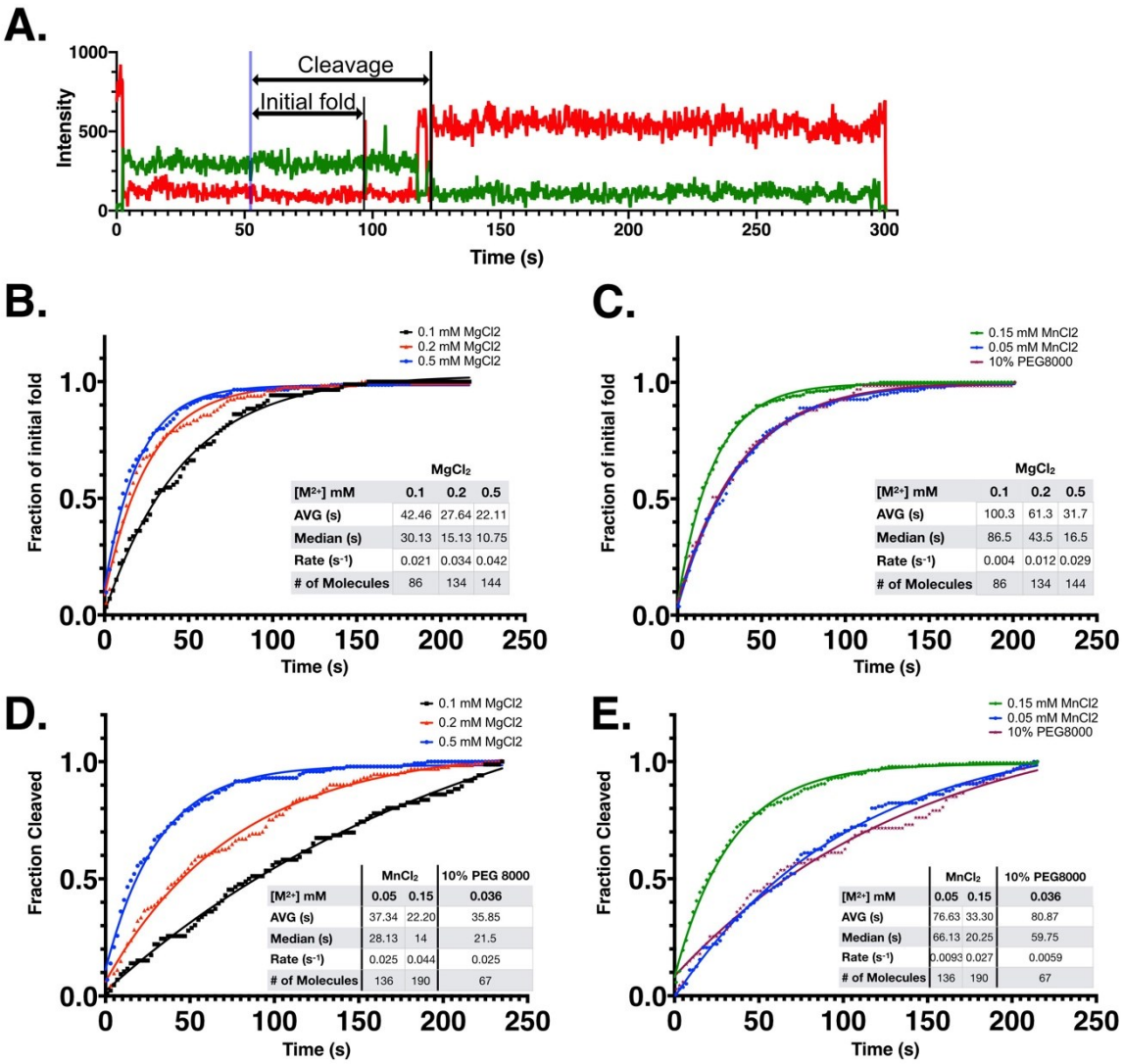
The midpoint of folding of the photocaged rU twister ribozyme is 0.37 ± 0.056 mM $MgCl_2$, with a Hill constant of 1 (Figure 12C). A large fraction of the rU variant did not respond to 405 nm light or were damaged and cannot adopt its tertiary fold. Instead of plateauing at 83% folded, as the non-photocaged dU twister ribozyme, it plateaued at 54% folded (Figure 12C). The same behavior can be observed in the photocaged dU twister ribozyme, which plateaued at 55% folded, but with a midpoint of 0.21 ± 0.084 mM $MgCl_2$ and a Hill constant of 1 (Figure 12B). The non-photocaged dU twister ribozyme has a midpoint of 0.16 ± 0.024 mM $MgCl_2$ and a Hill constant of 1, similar to the midpoint of cleavage seen in the gel based assay of 0.19 mM $MgCl_2$, and the photocaged dU twister ribozyme (Figure 12C). Although the photocage works in

a population of molecules, there is a second population that does not respond to activation or is damaged.

Rate of ribozyme folding and cleavage under different $MgCl_2$ concentrations

By measuring the rate of folding following photocage activation, I can ascertain the kinetics of the first folding events and characterize how folding is coupled with cleavage. To do this, I measured the rate of initial folding as the time spent in the unfolded state between photocage activation and the first folding event. Similarly, the rate of cleavage was measured as the time before the persistent high FRET state which is characterized as having a lifetime ≥ 30 s (Figure 14A). I specifically focused on molecules which achieved a persistent high FRET state within the length of the movie, because I could not differentiate molecules which failed to be deprotected from those that did not happen to fold within the duration of the movie. When I included all molecules that have at least one folding event in my analysis, the rates of folding became independent of $MgCl_2$ concentration, something that does not match what is seen in gel-based assays and single molecule folding rates. This $MgCl_2$ independence is due to a population of molecules not being activated by blue light exposure.

For the active (uncaged) molecules, I find that as $MgCl_2$ concentration increases, the rate of initial folding also increases, suggesting that the higher divalent ion concentration is stabilizing the folded state (Figure 14B). Similarly, the rate of cleavage increases with increasing $MgCl_2$ (Figure 14C). Interestingly, the average time until self-cleavage approaches the average first folding time at 0.5 mM $MgCl_2$ (Figure 14 C,D). This is consistent with the initial event becoming the cleavage event, as the delay between photocage activation and cleavage becomes shorter as observed at very high $MgCl_2$ concentrations, as seen in100 mM (Figure 12A).



MgCl₂			
[M²⁺] mM	0.1	0.2	0.5
AVG (s)	42.46	27.64	22.11
Median (s)	30.13	15.13	10.75
Rate (s⁻¹)	0.021	0.034	0.042
# of Molecules	86	134	144

MgCl₂			
[M²⁺] mM	0.1	0.2	0.5
AVG (s)	100.3	61.3	31.7
Median (s)	86.5	43.5	16.5
Rate (s⁻¹)	0.004	0.012	0.029
# of Molecules	86	134	144

MnCl₂		10% PEG 8000
[M²⁺] mM	0.05	0.15	0.036
AVG (s)	37.34	22.20	35.85
Median (s)	28.13	14	21.5
Rate (s⁻¹)	0.025	0.044	0.025
# of Molecules	136	190	67

MnCl₂		10% PEG8000
[M²⁺] mM	0.05	0.15	0.036
AVG (s)	76.63	33.30	80.87
Median (s)	66.13	20.25	59.75
Rate (s⁻¹)	0.0093	0.027	0.0059
# of Molecules	136	190	67

Figure 14: Folding conditions modulate ribozyme efficiency

A) An example trajectory from a photocaged rU twister showing the time of photocage activation with a blue vertical line. The initial folding time is defined as the interval from the end of 405 nm exposure to the start of the initial folding event. The time to cleavage is defined from the end of 405 nm exposure and the beginning of the persistent high FRET state, which represents the product complex.

B) The frequency distribution for the time to first folding event in 0.1, 0.2 and 0.5 mM $MgCl_2$. The cumulative distribution was fit to a single-phase exponential rate equation with a burst phase.

C) The frequency distribution of cleavage time, as in B.

D) Time to the initial folding event in 0.036 mM $MgCl_2$ + 10% PEG8000 and $MnCl_2$. 0.036 mM $MgCl_2$ + 10% PEG8000 is isostable with 0.1 mM $MgCl_2$, while 0.05 mM and 0.15 mM $MnCl_2$ are isostable with 0.1 and 0.5 mM $MgCl_2$ respectively.

E) Time to self-cleavage in PEG and $MnCl_2$ as in D.

Effect of $MnCl_2$ and PEG8000 on twister ribozyme folding

In order to understand the effects of $MnCl_2$ and PEG8000 on the initial folding rate and cleavage rate, isostable conditions for each variable were used. The isostable conditions were defined as the concentrations of $MnCl_2$ or PEG8000 under which the fractions of folded molecules are equal to those in $MgCl_2$. The fraction of folded dU twister ribozyme in 0.05 mM and 0.15 mM $MnCl_2$ matched that in 0.1 mM and 0.5 mM $MgCl_2$, while 0.032 $MgCl_2$ was used in the presence of 10% PEG8000 this yielded a similar fraction of folded RNA as 0.1 mM $MgCl_2$ (Figure 7C). The initial folding rate was similar between all three isostable conditions (Figure 14 D). All the conditions which match 0.1 mM $MgCl_2$ (0.05 mM $MnCl_2$ and 0.032 $MgCl_2$ + 10% PEG8000) have similar average times for the initial folding event (40 s), as well as a similar rate of folding (0.025 s^{-1}). However, the rate of cleavage is higher in the presence of both $MnCl_2$ and PEG8000 relative to the rate of cleavage in $MgCl_2$. When compared to a higher 0.5 mM $MgCl_2$ concentration, 0.15 mM $MnCl_2$ cleavage rate equals that of $MgCl_2$ at ~0.028 mM. The initial folding rate is unchanged between the different conditions, however, the cleavage rate increases for $MnCl_2$ at low concentrations but evens out at higher concentrations, explaining how Mn2+ preferentially stabilized the folded state of the dU twister ribozyme (Figure 7A,C). At higher concentration, $MgCl_2$ is able to compensate for misfolding, that $MnCl_2$ is able to do at lower concentrations.

I then went on to compare the overall folding rate for all the conditions tested, to see if there is any obvious outlier. The overall folding rate looks at all the folding events that occur in the lifetime of each molecule (Figure 15). All the conditions provide the expected folding rate, with the folding rate increasing with $MgCl_2$ concentrations, and the isostable folding rate being close to those of the their $MgCl_2$ counterparts (Table 1). Under equal $MgCl_2$ concentration, the

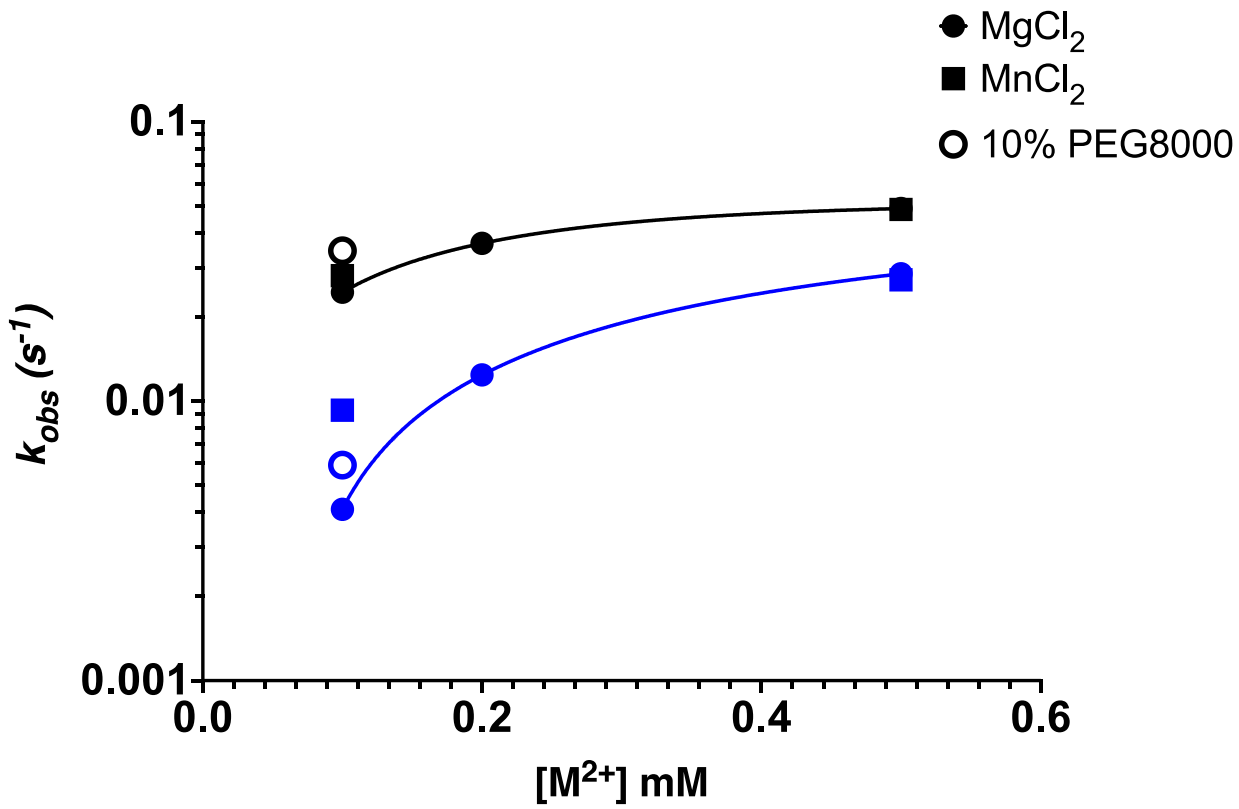


Figure 15: Comparing folding rate to cleavage rate under different ion concertation

The rates of folding and cleavage under different concentration of divalent metal ions. The folding rates are in black, while the cleavage rates are in blue. Circle represents $MgCl_2$, square represent $MnCl_2$ and open circles are $MgCl_2$ with 10% PEG8000. k_{obs} vs. $MgCl_2$ is fit to a single exponential. The folding rates in $MnCl_2$ and $MgCl_2$ plus 10% PEG 8000 are plotted at the isostable Mg^{2+} concentration.

Table 1: Calculated rates of folding after photocage activation under different isostable conditions.

Table 1: Folding Rates	**$MgCl_2$**			**$MnCl_2$ mM**		**10% PEG8000**
$[M^{2+}]$ mM	**0.1**	**0.2**	**0.5**	**0.05**	**0.15**	**0.036**
Initial Fold (s^{-1})	0.021 ±0.002	0.038 ±0.004	0.049 ±0.003	0.027 ±0.002	0.046 ±0.002	0.027 ±0.002
Folding Rate (s^{-1})	0.024 ±0.001	0.037 ±0.001	0.051 ±0.004	0.028 ±0.001	0.048 ±0.002	0.036 ±0.001
Folding Rate w/o Initial (s^{-1})	0.026 ±0.004	0.036 ±0.002	N/A	0.029 ±0.002	N/A	0.043 ±0.003
Cleavage (s^{-1})	0.0044 ±0.001	0.013 ±0.001	0.032 ±0.002	0.0096 ±0.001	0.030 ±0.001	0.0069 ±0.001
Number of Molecules	86	134	144	136	190	67

Initial fold is the dwell time of the low FRET (~0.1), unfolded molecule, from the time of activation with 405 nm light, to the first high FRET (~0.9), folded event. The folding rate is the combination of dwell times the twister ribozyme spends in the unfolded state. Folding rate without the initial fold, is the time the molecule spends in the unfolded state, without the initial high FRET test state. Cleavage is the time between 405 nm light activation and the long-lived high FRET state. Errors represent the 95% likelihood confidence interval calculated using GraphPad Prism 8.

initial folding rate, the overall folding rate, and the folding rate without the initial folding event are relatively equal. However, for high $MnCl_2$ concentration, the initial fold happens quicker by 0.01 s^{-1}. Meaning that once twister ribozyme performs an initial fold, its probability of folding again is lower than it was to perform the first fold. While $MgCl_2$ increases the initial folding rate and subsequent folding rates equally, $MnCl_2$ seems to increase the initial fold the most, while the other general folding rate lags behind, explaining how the cleavage rate could match at higher divalent ion concentrations.

PEG8000 is the opposite, with a slow initial fold (0.025±0.002 s^{-1}), and then quick folding rates after (0.043 ±0.003 s^{-1}). PEG8000 seems to stabilize a structure that needs some time to attempt to fold, but once it folds and unfolds, its probability of re-folding almost doubles. This can explain why PEG8000 has the biggest effect at lower $MgCl_2$ concentrations. At low $MgCl_2$ concentrations, the chance of a molecule folding is low for both the initial and subsequent folds, 0.021±0.002 and 0.026±0.004 s^{-1} respectively, however, in PEG8000, the initial fold is equals to that of $MgCl_2$ alone (0.021±0.002 s^{-1}), but its subsequent folding probability jumps once it stabilizes the folded, compact state. Therefore, PEG8000 seems to function by destabilizing the unfolded state, just as others have seen before (Kilburn et al. 2010).

Testing activation of the photocage on nucleotide G24

The photocage positioned at G7 had proven to be beneficial for understanding the dynamics of twister ribozyme. However, I wanted to see if the photocage transferred to a secondary position on the same pseudoknot functions similarly to G7PC (Figure 16A). To test this, I generated a twister ribozyme containing a photocage on nucleotide G24 which is located on the opposite side of the same pseudoknot as G7. G24 stabilizes the T2 photocage but does so from the opposite strand to G7.

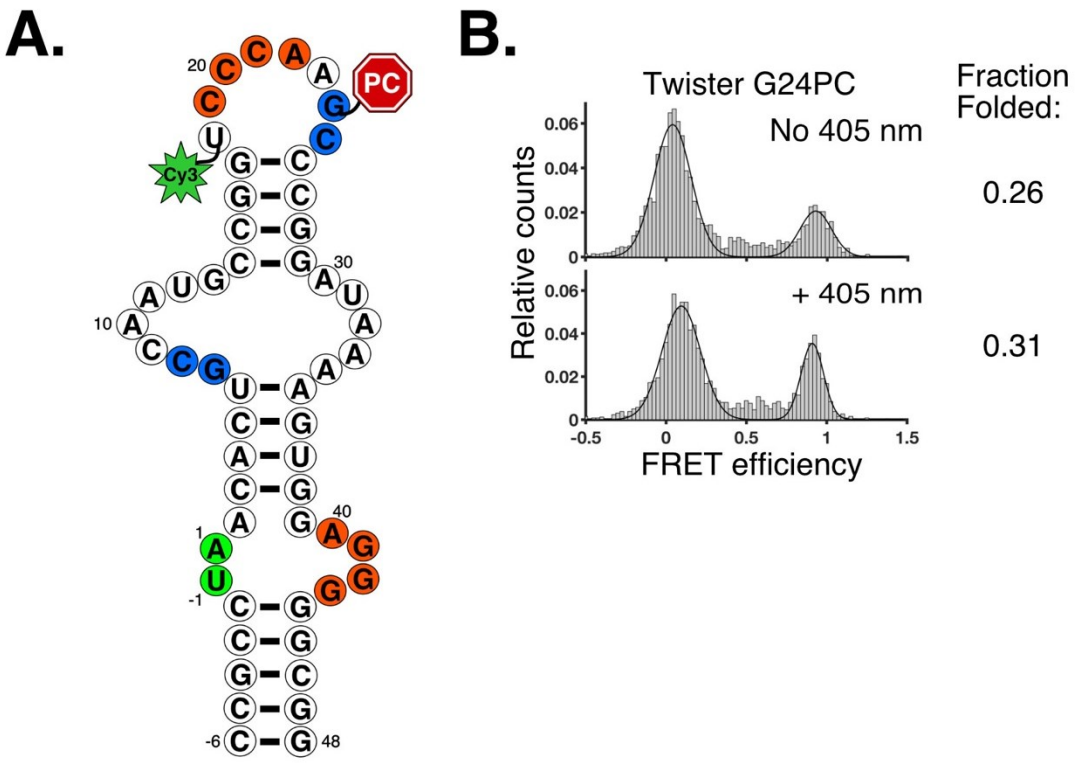


Figure 16: Activation of G24PC twister ribozyme with blue light

A) Schematic of twister ribozyme G24PC (red octagon), which has the photocage on the opposite side of pseudoknot 2.
B) The activation of the photocage at the G24 nucleotide position. A population has folded before exposure to 405 nm light. After 405 nm exposure, the folded population increases slightly from 26% to 31%.

I found that the G24PC twister ribozyme was still activated by blue light, but it did not activate to the same extent as G7PC twister ribozyme in 2 mM $MgCl_2$ (40%) (Figure 16B). It is also obvious that the photocage at G24 did not completely prevent folding, with 26% of the population already folded before blue light activation. With blue light activation, the folded population slightly increased to 31%. This discrepancy between G7PC and G24PC might be due to interactions between the photocage and its neighboring nucleotides. I also found that the photocage began to respond to the green laser at a wavelength of 532 nm. The response was not as fast as to the blue light, requiring seconds of exposure instead of microseconds to activate the photocage. During the purification process, the photocaged RNA is scanned for Cy3, meaning that it is exposed to a green laser, which could deprotect a population of the RNA, resulting in the 26% folded population before blue light activation.

Discussion

Twister ribozyme is a great candidate for RNA folding studies due to its short size and availability of crystal structures (Liu et al. 2014; Ren et al. 2014). Furthermore, its activity can be modulated using ribozymes that are modified with fluorophores and photocages to better explore its dynamics and self-cleavage. I developed a novel tool to successfully control twister ribozyme dynamics with photochemistry allowing me to decouple the secondary fold from the tertiary fold with a fast activating photocage. This also marks the first attempt at studying a photocaged RNA using TIRF microscopy.

Initial studies showed that the twister RNA can be modified with a fluorophore without disrupting the folded RNA a significant amount. Our lab's previous attempt to label twister ribozyme had proven somewhat flawed, as the fluorophore within the backbone disrupted the tertiary fold of twister and required a hundred times higher concentration of $MgCl_2$ to reach the

folding midpoint. The rates of cleavage calculated for U18 are similar to those that were published previously with related, unlabeled twister ribozymes, showing that the new labeling scheme did not disrupt the folded structure. The midpoint of cleavage observed in the gel-based cleavage assay (0.19 ± 0.019 mM), nearly matches the midpoint of folding we observed by smFRET spectroscopy (0.16 ± 0.021 mM), and is close to what was measured for the photocaged dU twister ribozyme (0.21 ± 0.083 mM $MgCl_2$). Compared to the midpoint of folding observed previous by Panja et al. using smFRET spectroscopy (15 mM $MgCl_2$) the fluorophore label, as well as the photocage, did not significantly disrupt the folding and subsequent cleavage of the ribozyme (Panja et al. 2017a).

The Greenberg lab, which was responsible for synthesizing the photocage and then incorporating it within the RNA, had succeeded in shifting the activation wavelength of the photocage into the visible spectrum (405 nm). The experiments discussed here showed that the photocage is active when being incorporated into an RNA, responds to 405 nm light, is fast activating and does not disrupt the structure of the unfolded or folded states. Our time activation limit was 0.03, due to our shutter and camera speed, which is the shortest activation time we tested in the smFRET experiments. However, this photocage hypothesized to be removed within the μs timescales, which would require different instrumentation to test (Klán et al. 2013). I found that this technique could be multiplexed and performed in a high-throughput manner, since different fields of view can be activated individually allowing for multiple experiments to be performed in the same channel, without needing to do flow-in experiments, as was previously required.

When incorporated on the G7 nucleotide of the twister ribozyme, the photocage succeeded in protecting the ribozyme in the unfolded state. Looking at the population

histograms, it becomes clear that the photocage is protecting the unfolded state, through a range of concentration of $MgCl_2$ (Figure 12B). The unfolded state has the same FRET values as seen previously, with the same being true for the folded state, highlighting that the we are looking at the same structures that occur without the photocage. Furthermore, the rates of folding observed at 0.1 and 0.2 mM $MgCl_2$ in the photocaged rU twister ribozyme (k_F = 0.025 ±0.001 and k_F = 0.037 ±0.001) matches the rates of folding of the non-photocaged dU twister ribozyme (k_F = 0.022 ±0.002 and k_F = 0.03 ±0.0002). This suggests to us that the structure of the activated ribozyme is not affected and the dynamics we witness after photocage activation are biologically relevant.

The rate of the initial fold and the rate of folding match throughout MgCl2 concentrations (Table 1). This highlights that the photocage does not disrupt the secondary fold of the twister ribozyme, because it does not have a slower rate of folding right after 405 nm light activation. The rates are equal between the initial fold and general fold, meaning that the photocage does not confine the twister ribozyme into an unproductive structure.

In the presence of high $MgCl_2$ concentration, the G7PC ribozyme samples the folded state (Figure 12B) indicating that this structure may form but is highly unstable when the photocage is present. The appearance of these transitions at high $MgCl_2$ concentration could be due to Mg-dependent compaction folded state possibly by supporting partial formation of the main pseudoknot labeled in orange (Figure 11A). However, due to the photocage, the secondary pseudoknot is not available to lock the structure into the folded state, resulting in rapid unfolding. The photocaged rU twister ribozyme had a midpoint of folding of 0.37 ± 0.16 mM $MgCl_2$, hinting at the disruption of its structure via the photocage. However, I believe this is not the case, as its more likely that there is a population of photocaged molecules that are not

responding to 405 nm light. At low $MgCl_2$ concentrations, there are a few molecules that fold with and without the photocage. However, as the $MgCl_2$ concentration increases, the fraction of folded molecules is being suppressed by molecules which are unable to be activated by 405 nm light. The same effect can be seen with the photocaged G24, whose activation by 405 nm light is greatly suppressed. There is a large possibility that the photocage interacts with the RNA structure, which prevents it from being activated by blue light. This same structure formation can also explain how G24 is able to be activated by a 532 nm green laser. Therefore, when designing a photocaged RNA, the surrounding nucleotides and structure should be taken into account in order to see a proper benefit.

Testing the effect on folding rates at different concentrations of $MgCl_2$ within the transitional phase allowed me to detect unproductive folding of twister ribozyme, before its final cleavage. At low $MgCl_2$ concentrations, there are several folding attempts before persistent folding and cleavage is achieved. The number of these initial folding attempts decreases as $MgCl_2$ concentration increases. Furthermore, as the concentration of $MgCl_2$ increased, the average time until both the initial folding event and cleavage decreases (Figure 15). At some point, the initial folding rate and the cleavage rate will intersect as the molecule would fold into its persistent high FRET state right after activation. These experiments cannot detect the nucleotide level rearrangements that might occur within the active site required for cleavage, only the formation of stable pseudoknots can be detected. At lower $MgCl_2$, the active site rearrangement might require the molecule to unfold, and refold until it cleaves, stabilizing the folded state. However, at high $MgCl_2$ concentration, the rearrangement can be done in the folded state, which is stabilized by increased metal ion concentration. Therefore, cleavage might occur

later in the lifetime of the folded state but due to our fluorophore positioning, we can only detect the persistent state which we consider to be the folded state.

It has been well established that twister ribozyme has a lower midpoint of folding in the presence of Mn^{2+} and that PEG8000 increases the population of the folded state at the same $MgCl_2$ (Panja et al. 2017a). Testing isostable conditions in the presence of $MnCl_2$ and PEG8000 was done to understand the reasons for these observations. At first, I hypothesized that these conditions will allow me to detect and stabilize intermediates that occur during folding, by watching the dynamics at these conditions. However, no stabilized intermediates were detected in either $MnCl_2$ or PEG8000 conditions. This is likely due to the time resolution limit of 30 ms. Previous labs were able to detect intermediate steps in different ribozymes only after increasing their resolution of 10 ms (Tan et al. 2003). I observed that $MnCl_2$ affects the twister ribozyme by increasing its initial folding rate, while keeping the general folding rate relatively equal, as the $MnCl_2$ concentration is increased. This seems to point to $MnCl_2$ destabilizing the unfolded state and favoring formation of the folded state which in turn increases the chances of cleavage. The opposite may be true in the presence of 10% PEG8000, where it increases the general rate of folding, while keeping the rate of the initial fold low. This explains the effect of PEG8000 at low $MgCl_2$ concentrations. The initial rate of folding is slow, once it folds and unfolds, the chances of it refolding doubles, resulting in a larger population of folded molecules at lower $MgCl_2$.

RNA structural formation is critical for function, therefore studying RNA structural dynamics and folding rates is instrumental in understanding function. Creating a photo-activatable self-cleaving ribozyme opens many avenues for the study of RNA folding and cleavage under different environmental conditions. By creating a photo-activatable RNA, the experimenter becomes the conductor of the RNA, deciding when the RNA should undertake its

journey though the folding pathway. Here we develop and test a fast acting photocage on the twister ribozyme to better understand its unfolded to folded transition and cleavage.

Materials and Methods:

RNA synthesis

All RNA sequences were based on the *Oryza sativa* twister ribozyme, 5'-r(CCGCCUAAC ACUGCCAAUGCCGGUCCCAAGCCCGGAUAAAAGUGGAGGGGGCGG), with a 38 nt SA5 DNA extension at the 3' end d(AGGACGACACACTTTGGACAGGACACACAGGAC ACAGG), resulting in a 92 nt long *Ost* Twister-SA5 RNA. All sequences either order through Invitrogen, IDT or custom made for this study were gel purified before being used in ligation reactions. Variations of Cy3-labeled *Ost* Twister-SA5 ribozyme were prepared by enzymatic ligation of RNA fragments (Table 2) corresponding to the 5′ half of twister, 5′- Ost Twister r(CCGCCUAACACUGCCAAUG CCGGU$^{2'Cy3}$CCCA), and the extended 3′ half of twister, 3′Ost Twister-SA5, with the aid of a DNA splint, OSTSplint. Various 5′ *Ost* Twister fragments were prepared as follows: The 5′- *Ost* Twister self-cleaving RNA with a photocage at position G7 was made by ligating 5′OST11nt and 5′-3′OstG7PCU18Cy3 using 5'OST-LigSplint. 5′-3′OstG7PCU18Cy3 was synthesized in-house using standard phosphoramidites, 2'amino-uridyl phosphoramidite (Glen Research) for attachment of Cy3, and the custom synthesized photocaged G phosphoramidite (G^{PC}) (Huabing Sun). A photocaged non-self-cleaving variant containing 2′-dU at position 6 was prepared in the same way using 5'OST11ntdU. The non-cleavable, non-photocaged twister RNA was prepared by ligation of 5′Ost-dU-U18cy3 and 3′Ost Twister-SA5 without a splint. G24PC was generated by ligating 3′OstG24PC, with the rest of the SA5_-4 sequence, suing SA5 splint, to create a photocaged 3′Ost Twister-SA5. Then the 3′Ost G24PC Twister-SA5 was ligated to 5′- Ost-cy3, to make the full-length Ost twister ribozyme

Table 2: RNA and DNA oligonucleotides used in this study

	Name	Sequence	Source
1	3′Ost Twister-SA5	r(AGCCCGGAUAAAAGUGGAGGGGGCGG)d(AGGACGACACACTTTGGACAGGACACACAGGACACAGG)	IDT
2	5′- Ost-Cy3	r(CCGCCUAACACUGCCAAUG CCGGU$^{2'Cy3}$CCCA)	this study
3	OSTSplint	d(CCGCCCCCTCCACTTTTATCCGGAACCGG CAGGCAGTGTTAGGCGG)	Invitrogen
4	5′OST11nt	r(CCGCCUAACAC)	IDT
5	5′-3′OstG7PCU18Cy3	r(UG^{PC}CCAAUGCCGGU$^{2'NH2}$CCCA)	this study
6	5'OST11ntdU	r(CCGCC)d(U)r(AACAC)	this study
7	5′Ost-dU-U18cy3	r(CCGCC)d(U)r(AACACUGCCAAUG CCGGU$^{2'Cy3}$CCCA)	this study
	5'OST-LigSplint	d(TGGGACCGGCATTGTTTAGGCGG)	IDT
8	3′OstG24PC	r(AGCCCGGAUAAAAGUGGAGGGGGCGG)d(AGGAC)	this study
9	SA5_-4	d(GACACACTTTGGACAGGACACACAGGAC ACAGG)	Invitrogen
10	SA5 Splint	d(GTGTCCTGTCCAAAGTGTGTCGTCCGCCC TCCACTTTTATCC)	Invitrogen

All RNA oligonucleotides were gel purified on a denaturing 16% polyacrylamide gel.

U18Cy3G24PC. Biotin-SA5-Cy5 was used to anneal to SA5 containing Twister molecules and tether them to quartz slides. All procedures involving the photocaged RNA were done in low light to prevent the photocage from being activated.

Fluorophore labeling

Before ligation, amine modified RNAs were labelled with sulfonated Cy3 (GE Amersham Cy3 or Cy5 Mono-Reactive Dye Packs, catalog number PA23001 and PA25001), respectively. 5 nmols of RNA was combined with 33 µl of entirely resuspended fluorophore in DMSO and the volume brought up to 100 µl with 100 mM $NaCO_3$, pH 8.6 in a dark tube and left overnight at room temperature (~20°C). The reaction was then divided into two 50 µl aliquots and passed over two CHROMA SPIN+TE10 columns (Takara, cat# 636066) to separate the free dye from the RNA. The flow-through from each column was combined and the RNA precipitated with 3 volumes of ethanol and 15 µg glycogen and incubated at -80 °C for one hour. The RNA precipitate was collected at 14,000 rpm for 30 min, washed with 70% ethanol, and again spun at 14,000 rpm for 30 min. The liquid was decanted, and the RNA dried under vacuum in a SpeedVac SC100. The RNA was dissolved in 10 µl deionized water (18.2 MΩ, RNase-free) and stored at –20 °C.

RNA ligation

All ligations were done using the method of Stark and Rader with some modifications (Stark and Rader 2014). The 3′ RNA (12 µM final) was phosphorylated with 40 U T4 PNK (NEB) in the presence of 1.5 mM ATP and 12 U RNasin Plus (Promega) in a total volume of 40 µl at 37 °C for 30 min, and then denatured at 65 °C for 10 min (Figure 9). The reaction was placed on ice and the 5′ RNA fragment was added, with or without the DNA splint, in the ratio of 1:1.2:2 RNA1:Splint:RNA2 (Rodgers et al. 2016). The volume was brought to 46 µl with

deionized water and annealed in a thermocycler by heating to 90 ℃ for 2 min and decreasing the temperature by 5 ℃ every 2 min, until the temperature reached 5 ℃. The reaction was then placed on ice and 20 U T4 RNA ligase 1 (NEB) was added in the presence of 1X ligation buffer (NEB) and another addition of 1.5 mM ATP, and incubated 16 hrs at 16 ℃. The RNA products were resolved on a denaturing 16% (28 nt RNA) or 10% (92 nt RNA) polyacrylamide gel. The bands were scanned for a Cy3 signal on a GE Typhoon, alongside a control RNA of similar in length as a size marker. The desired product bands were excised, and the RNA extracted by freeze-thaw and soaking for 16 hrs in 1X TE Buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA). The RNA was then precipitated with 3X volume of ethanol, and spun at 9K RPM, 4 ℃ for 30 mins. The supernatant was discarded, and the pellet washed with cold 70% ethanol, and again spun down at 9K RPM, 4 ℃ for 30 mins. The pellet was then resuspended in 10 µl of deionized water.

Twister activity assays

All activity assays were done as previously described (Panja et al. 2017a). A mixture of 40 nM Ost-Twister and 60 nM of Biotin-SA5 DNA in 1X HK buffer (30 mM K-HEPES, pH 7.5; 100 mM KCl,) was incubated at 75 ℃ for 5 min, followed by 37 ℃ for 15 min, and 25 ℃ for 5 min. The mixture was then kept at room temperature and 4 µl were added to PCR tubes containing 1 µl 5X $MgCl_2$. After 1 min, the reaction was quenched with 2X TBE urea loading buffer (8 M urea, 90 mM Tris-HCl, 89 mM Boric Acid, 0.04% bromophenol blue and xylene cyanol, with 50 mM EDTA). To detect the cleavage of Twister over time, the annealed RNA was added to 5X $MgCl_2$ (2 mM final) and the reaction was quenched at various times with the addition of 2X TBE denaturing loading dye. The reaction products were resolved on a denaturing 10% polyacrylamide gel for 2 h at 15 W. The gels were scanned for the Cy3 signal (GE

Typhoon). The bands were quantified using ImageQuant 5.2, and the relative fraction cleaved was plotted in Graphpad Prism 8.

In order to check for removal of the photocage protecting group in bulk samples, 2 µl of the photocaged Twister mixture mentioned above (2 µl 40 nM twister•biotin-SA5) was exposed to 395-405 nm blue LEDs in a custom-made dark box. A motorized sample holder was used to rotate the sample as it was exposed for different lengths of time. After each exposure, the tube was transferred to a dark container. Then the exposed twister solution was added to tubes containing $MgCl_2$ and allowed to react for 1 min at room temperature. The reaction was quenched with 2X TBE denaturing loading dye and the products resolved on a denaturing 10% polyacrylamide gel for 2 h at 15 W. The gel was scanned and analyzed as described above.

Single molecule FRET assays

In a 10 µl reaction, 40 nM Twister RNA was combined with 20 nM Biotin-SA5-Cy3 DNA in 1X HK Buffer with 12 U RNasin Plus. The reaction was incubated at 75 ℃ for 5 min, 37 ℃ for 15 min, 25 ℃ for 5 min, and then kept at room temperature. The mixture was diluted 700-fold in 1X HK buffer and loaded right away onto a covered quartz slide passivated with dichlorodimethylsilane (DDS, Sigma, cat# 440272-100ml) 0.2 mg/ml biotin-BSA (Sigma cat# A8549-10MG), 2% Tween 20, and 0.2 mg/ml streptavidin as previously described (Hua et al. 2014). The slide was incubated with this mixture for 1 min, before the unbound RNA was washed away with imaging buffer containing 1X HK buffer, $MgCl_2$ or $MnCl_2$, 0.8% glucose, 4 mM Trolox, 0.1 mg/mL glucose oxidase and 0.02 mg/mL catalase. Reactions also contained 10% PEG8000 where stated.

The data were acquired using a home built prism-TIRF microscope, composed of an Olympus IX73 microscope with an Andor iXon Ultra EMCCD camera, CrystalLaser 405 nm

blue laser (19 mW output), 633 nm red laser (19 mW), and a 532 nm green laser (63 mW) (Roy, Hohng, and Ha 2008). Intensities in the green and red channels were recorded with 532 nm excitation. Each movie started and finished with ten frames of 633 nm excitation to detect the presence of the acceptor. All population histograms were built using 40 frame long movies, with each frame being 50 ms. At least 10 movies, with more than 200 molecules were used to build each histogram. Histograms showing the population after exposure to the 405 nm laser for 250 ms or 30 ms, were recorded 2 min after exposure to the blue laser, after twister molecules reach equilibrium. Long movies involving 50 ms frames were recorded for 1020 frames, or ~ 1 min, while movies used for the detection of first folding and cleavage events were taken for ~ 5 min with 250 ms frames, with 405 nm laser exposure occurring at 52.5 s for 1 frame.

Single molecule data analysis

Trajectories showing the intensities of single molecules over time were extracted from the raw movies using custom IDL code (Hua et al. 2014, 2018). Only molecules that have a Cy5 signal in the first ten frames of a movie, which were used for direct excitation correction, and had a Cy5 and a Cy3 signal within the first twenty seconds, which were used for leakage corrections, were selected for further analysis.

Population histograms

Population histograms were generated from a combination of ten movies, split into three batches, composed of 40 frames, with the middle 20 frames used to show the state of the molecule (FRET state). In making the histograms, molecules were not selected in any way other than having both Cy5 and Cy3 active fluorophores. Corrections for the background, reflection, leakage and direct excitation were made using the principals published previously (Hellenkamp et al. 2018). Histograms of the FRET efficiency were fit to two Gaussian equations (a1*exp(-((x-

b1)/c1)^2) (MatLab) to calculate the area under the curve for the high FRET state and the low FRET state. The ratio of folded molecules to total molecules, f_F = (area highFRET / area highFRET + area lowFRET) was calculated for each third of the data. These three values were used to calculate a mean and standard deviation for f_F at each condition. The change in f_F with divalent metal ion concentration, $[M^{2+}]$, was fit to a Hill equation $f_F = [M^{2+}]n/([M^{2+}]_{1/2}{}^n + [M^{2+}]^n)$and the midpoint ($[M^{2+}]_{1/2}$) of folding and hill constant (h) extracted. $MgCl_2$, $MnCl_2$ or PEG8000 concentration versus relative fraction folded were graphed and analyzed using Prism 8 graphing software.

Rates of folding and unfolding

To get the rate constants of folding or unfolding, 5 min movies of non-cleavable dU twister ribozyme were recorded with different $MgCl_2$ concentrations. Each trace was selected by hand, throwing out any molecules with quick bleaching event (within the first minute of the movie) or missing an active Cy3 or Cy5 fluorophore. The Cy3 and Cy5 intensities were then corrected for background, reflection, leakage and direct excitation and traces were selected when the molecule had more than one folding or unfolding event. Hidden Markov Modeling was done to get rates of folding and unfolding of non-cleavable twister ribozyme as previously described (McKinney, Joo, and Ha 2006).

Rates for initial folding and cleavage

Custom MatLab codes, as those used above, written and kindly provided by Boyang Hua from the Ha lab, were used to sift through 5 min long traces for those which had stable and active Cy3 and Cy5 fluorophores, which were selected for further analysis. Traces were analyzed by hand and the time from 405 nm laser exposure (52.5 s after the start of the movie) until the first folding event and subsequent folding events were analyzed within MatLab by selecting the start

and end times of the highFRET state. Start of high FRET was identified as the first instance when the Cy5 signal increased, while the Cy3 signal decreases. The time from 405 nm exposure to the first folding event was labelled as the initial fold, and the last stable folding event that was longer than thirty seconds in the cleavable twister ribozyme is self-cleavage. Molecules that had one active fluorophore or fluorophores that bleaches quickly (within less than 1 min of the start of recording) were discarded. The molecules were discarded because their quick bleaching prevented them from being useful in calculating folding or cleavage rates. For non-cleavable dU twister ribozyme, the average times to the initial folding event or to self-cleavage were calculated only with molecules which showed at least one folding event (high FRET state) after 405 nm exposure. For self-cleaving (rU) twister ribozyme, molecules that ended with a persistent high FRET state were used to analyze the initial folding rate, general folding rate and cleavage rate. A frequency distribution histogram was built of the dwell times for folding and cleavage events, and then fitted to a one phase decay equation, to get the rates of initial folding, self-cleavage and folding rate of the molecule at different $MgCl_2$ concentrations. All rate calculations and plots were generated using GraphPad Prism 8.

Chapter 3: 30S Subunit Helix 1 Folding

Introduction: Studying the formation of Alternative Helix 1

Helix 1 within the 30S ribosomal subunit contains the first nucleotide sequences to be synthesized, but it is one of the last to fold. Previous data published has hinted at a possible competing structure that forms and refolds later during the maturation of the 30S subunit. This refolding event is thought to be done by a ribosomal assembly factor which can bind and stabilize helix 1 over the competing helix. In order to test this hypothesis, the formation of the competing helix was simulated using two annealed RNAs and challenged with proteins which are hypothesized to disrupt it and refold into the native helix 1. Studying this behavior can further our understanding of the type of interactions that occur between proteins and RNA, while providing more detailed information about the maturation of the 30S subunit.

Mutation in Helix 1 influences 30S subunit Maturation

Previous experiments have shown that upstream regions of the 16S rRNA affect 30S subunit maturation by acting in a folding catalyst like manner (Balzer and Wagner 1998). Mutations in the 16S rRNA itself has also been found to influence the maturation of the 16S rRNA. The Noller lab identified a mutation within helix 1 (H1), C23U, which results in cold sensitivity, increase in free 30S and 50S subunits, incomplete processing of 17S rRNA, and a decrease in the amount of 70S ribosomes. These are all phenotypes of improperly matured ribosomal subunits (Dammel and Noller 1993). This mutation destabilizes H1 by changing a G-C Watson and Crick base pair to a G-U (Figure 18), thereby increasing the free energy of H1 and making it less likely to form. If H1 does not form, then helix 2 (H2) cannot properly fold, which forms the active site for the 30S subunit (Figure 17). The central H2 pseudoknot is positioned at the core of the 30S subunit, bringing together both the 5′, central and 3′ domains of 16S rRNA

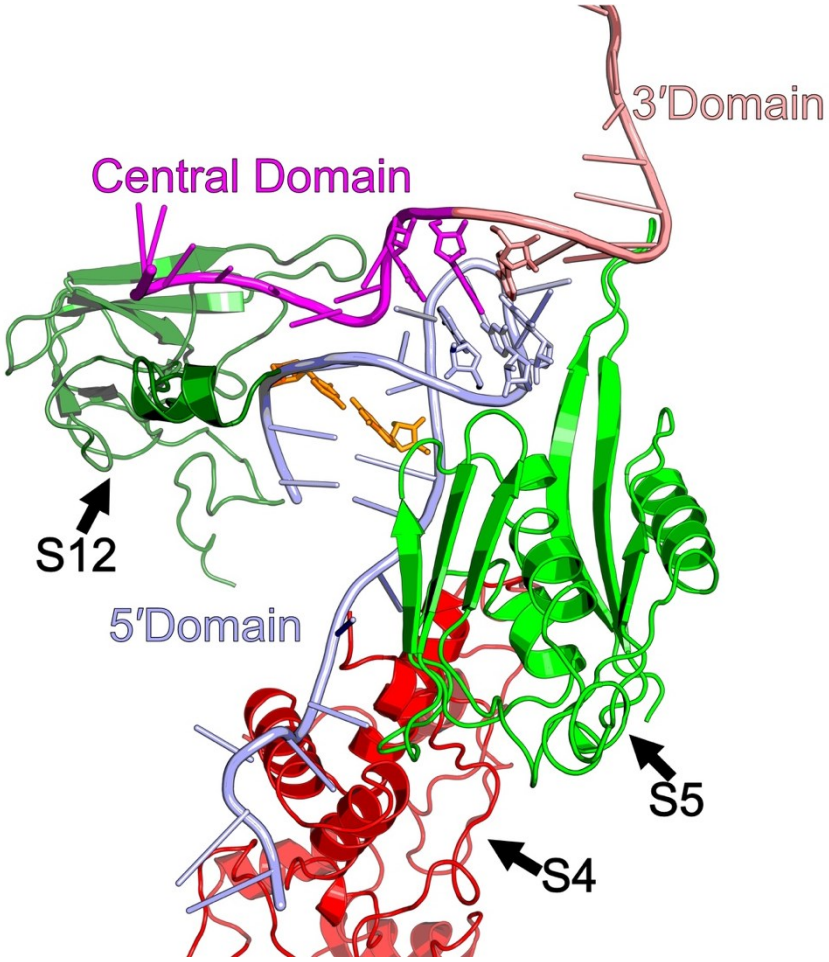


Figure 17: Proteins binding close to the central pseudoknot (Helix2)

The central pseudoknot brings together the 5′ domain in light blue, central domain in purple and 3′domain in light pink (PDB: 4V7S) (Dunkle et al. 2010). The disruption of base pairing due to the C23U mutation in helix 1 is highlighted in orange. Ribosomal proteins which bind close to the central pseudoknot and helix 1 are shown. S4 in red, signifying it's a primary binding protein, while S5 and S12 are different shades of green signifying they are secondary binding proteins. Mutations in S5 have cold-sensitive phenotypes linked with delayed 30S maturation. Protein S12 is a late assembly protein, and was identified as an interacting partner for Hfq, which is proposed to be a 30S assembly factor. S4 nucleates the folding of the 5′ domain and interacts with H1 in the mature 30S ribosome.

together (Figure 2). By disrupting or slowing down the formation of H2, the translation ability of the 30S subunit becomes disrupted (Brink, Verbeet, and Boer 1993; Poot et al. 1998).

Making compensatory mutations can rescue the cold sensitivity of the C23U mutation (Dammel and Noller 1993). Mutating G11A, rescued the Watson-Crick base pairing within H1, making H1 more favorable (Figure 18). The Noller lab also identified a mutation in the 5′LS, C-5U, that rescues the cold sensitivity of C23U. With this revelation, it was hypothesized that parts of the 5′LS make contact with 16S rRNA during folding, and a potential alternative-helix 1 (alt-H1) could form during maturation. Mutating C-5U destabilizes the hypothesized Alt-H1, making H1 more favorable in the C23U mutation. Overexpression of the assembly factor RbfA was also found to rescue the C23U phenotype suggesting that binding of RbfA may help overcome destabilization of H1 by C23U (Dammel and Noller 1995). RbfA is thought to bind in the late stages in maturation and promote processing of 17S to 16S rRNA, but the mechanism for how RbfA aids in maturation is not clear. These genetic studies suggest RbfA may function during formation of H1, H2, and the central pseudoknot perhaps by binding directly to H1 and stabilizing it (Figure 18). Although multiple labs have tried to visualize pre-30S particles using X-ray crystallography, electron microscopy, and footprinting, none have been able to identify the position of the 5′LS and the hypothesized alt-H1 (Yang et al. 2014). It is not known concretely if alt-H1 is capable of forming, where it would be located on the pre-30S, which parts of the 5′LS participate in the folding of 30S subunits, and what is the switch that signals the 5′LS to move from its chaperone function for cleavage. I want to detect the switching between alt-H1 and H1, and understand the molecular mechanism which plays a role in the switching of the helices. Using biochemical assays, I aim to show which proteins, if any, play a role in the stabilization of H1.

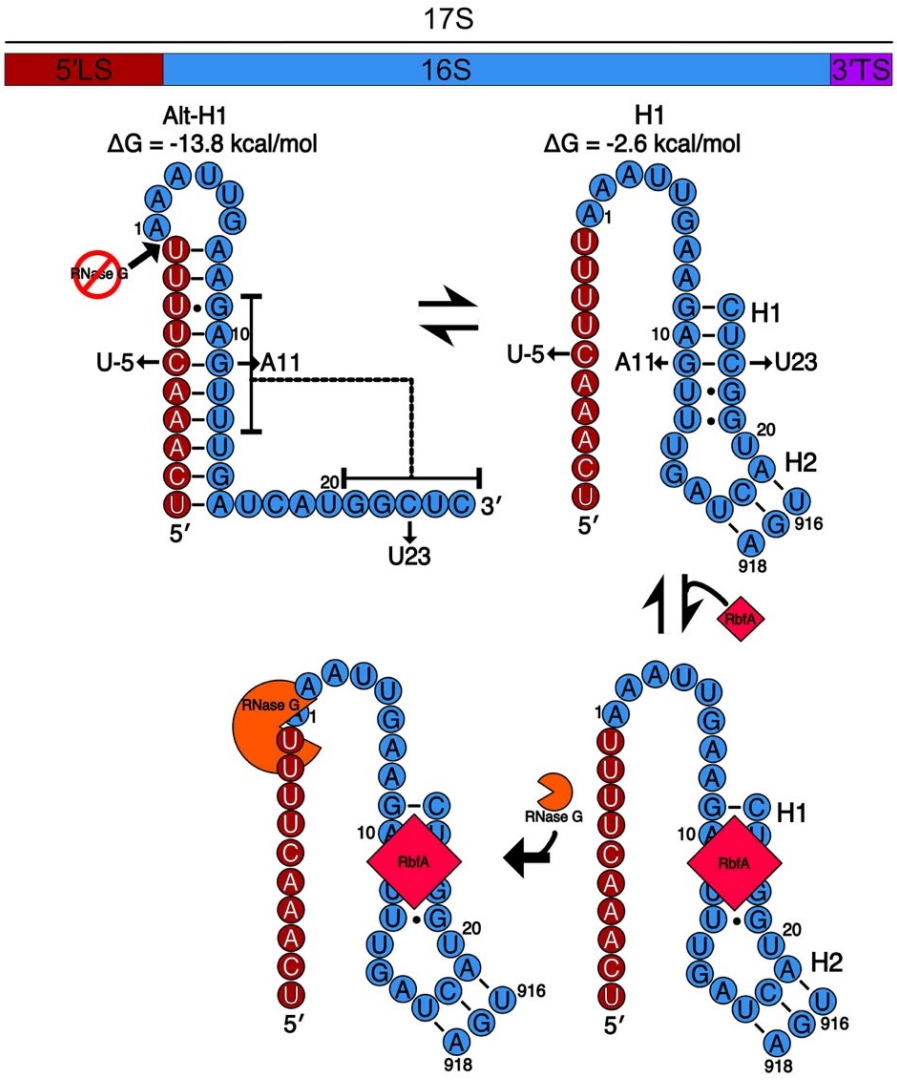


Figure 18: Competition between Alt-H1 and H1

It is hypothesized that part of the 5′LS, in red, makes an ALT-H1 (top left) by binding to the 5′ domain of 16S, in blue, sequestering the necessary nucleotides that form H1 (top right). Performing modeling calculations, it can be seen that the Alt-H1 has a lower free energy than H1, making more stable than H1. Making mutations that disrupt H1 (C23U), makes H1 less likely to form. However, compensatory mutations, such as G11A and could rescue H1 by restoring Watson-Crick base pairing (Dammel and Noller 1993). Cold-sensitivity can also be rescue by C-5U mutation which disrupts Alt-H1, making Alt-H1 unfavorable. Overexpression of the assembly factor RbfA (in pink), was found to rescue the cold-sensitive phenotype of the C23U mutant (Dammel and Noller 1995). It is hypothesized that RbfA is able to stabilize H1, allowing for RNase G, in orange, to come in and cleave the single-stranded RNA. This results in stabilizing H1 and the central pseudoknot at the late stages of 30S maturation. Figure derived from (Dammel and Noller 1993).

Results

5′LS forms a complex with 16S in trans

To study the competition between helix 1 and alternative helix 1 (alt-H1), I simulated alt-H1 formation by forming a complex in trans between native 16S rRNA and either the full length 5′LS (115 nts) or truncated 5′LS variants (Table 3). Initially, I used native gel mobility shift assays to resolve bound 5′LS•16S complexes from free 5′LS. The 115 nt WT 5′LS is capable of binding to the 16S rRNA at 42 ℃ and after a 65 ℃ incubation (Figure 19A, B). Because more 5′LS•16S complexes were present after 65 ℃ incubation, that was the temperature used for all the follow up experiments. Using 15 nt short 5′LS oligonucleotides, which are just long enough to form the predicted alt-H1 without any extra upstream sequence, I tested their ability to bind to 16S rRNA and form the alt-H1. I generated two mutant 5′LS oligonucleotides that are predicted to form more or less stable alt-H1 interactions. The 'ideal' mutant 5′LS oligonucleotide contains a U-3C mutation which converts a G-U wobble into a G-C base pair, creating a more stable alt-H1 helix (Table 3). I observed that both the WT and ideal 5′LS oligonucleotides could form a complex with the 16S rRNA. I also tested the C-5U mutation (labeled 'mutant' in Figure 19) that was previously identified by the Noller lab to disrupt alt-H1. This mutant 5′LS oligomer is unable to form a complex with the 16S RNA, suggesting that stable base-pairing is required for binding of the 5′LS to 16S RNA (Figure 19A, B) (Dammel and Noller 1993). These data suggest that the complex formed between the WT 5′LS and the 16S rRNA likely contains alt-H1 and is sequence specific.

In order to perform fluorescence anisotropy experiments, to detect the formation of the 5′LS•16S, fluorophore (FAM) labelled oligos were created. Although the anisotropy experiments did not take off, the FAM labelled sequences provided cleaner data for the native gel mobility

Table 3: 5′LS variants tested for trans binding to 16S

Name	Sequence	Length (nt)
5′LS	UGUGGGCACUCGAAGAUACGGAUUCUUAACGUCGCAAG ACGAAAAAUGAAUACCAAGUCUCAAGAGUGAACACGUA AUUCAUUACGAAGUUUAAUUCUUUGAGCGUCAAACUUU U	115
WT	GAGCGUCAAACUUUU	15
Ideal	GAGCGUCAAACUCUU	15
Mutant	GAGCGUUAAAUUCUU	15
FAM-5′LS	UGUGGGCACUCGAAGAUACGGAUUCUUAACGUCGCAAG ACGAAAAAUGAAUACCAAGUCUCAAGAGUGAACACGUA AUUCAUUACGAAGUUUAAUUCUUUGAGCGUCAAACUUU U-FAM	115
FAM-Ideal	UUUGAGCGUCAAACUCUU/3AmMO/	18
FAM-WT	UUUGAGCGUCAAACUUUU/3AmMO/	18
FAM-Mutant	UUUGAGCGUCAAAUUCUU/3AmMO/	18

Names and sequences of the different variants of the 5′LS used for intermolecular binding to 16S rRNA. The mutations are labelled in red. All the short oligos were received from IDT. 5′LS was transcribed from amplified PCR products and then labelled with ^{32}P. While FAM-5′LS was made through the ligation of an 18 nt RNA labelled with FAM with *in vitro* transcribed 97 nt RNA.

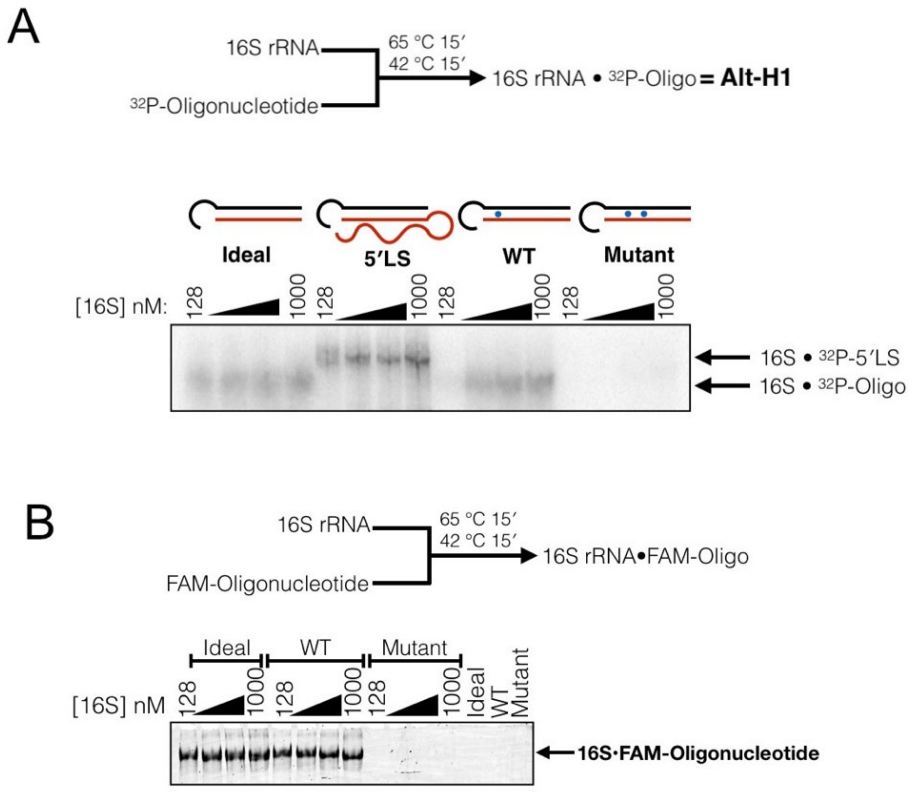


Figure 19: Binding of 5′ leader to 16S rRNA

A) The binding of ^{32}P-labelled 5′LS and its variants to 16S rRNA. 5′LS and its variants are colored in red, while 16S rRNA is black with blue dots representing G-U wobble base pairs. 5′LS, full-length 115 nt leader and WT, oligonucleotide that corresponds to the last 15 nt of the leader, both bind to the 16S rRNA. The fully matched (U–3C) Ideal oligonucleotide is also able to bind the 16S rRNA, but the Mutant (U–5C) oligonucleotide does not bind 16S rRNA. Top schematic shows how RNA samples were annealed before loading on the gel.

B) Similar to A, except that the oligonucleotides are labelled with a 6-FAM fluorophore. Once again, the Ideal and WT binding whereas the U-5C Mutant does not bind the 16S rRNA. The binding between 5′LS and 16S rRNA is sequence dependent.

shift assays. I repeated these gel shift binding experiments with fluorophore labeled oligonucleotides to confirm the results obtained with radiolabeled RNAs (Figure 19B).

Mature 30S subunits cannot bind 5′LS

Since the 5′LS binds to native 16S rRNA, I wanted to see if this would hold true for mature 30S subunits. Because the 30S subunit is matured, the central pseudoknot and H1 are folded, and r-proteins are bound to the rRNA preventing H1 from unfolding. I combined the 5′LS RNA with 30S subunits and incubated the mixture at 42℃, the same temperature used for 30S reconstitution, or at 65℃, knowing that the higher temperature would denature some of the r-proteins. 5′LS RNA cannot bind to mature 30S subunits (Figure 20A, lane 7), showing that H1 is stably folded and cannot be converted to Alt-H1 in the presence of the r-proteins. However, if the 30S subunit is denatured at high temperatures, 5′LS can bind to the partially denatured 16S rRNA (Figure 20A, lane 8). Because 5′LS clearly binds to naked RNA, I hypothesized that a rearrangement of the 16S rRNA during 30S subunit assembly unwinds alt-H1 and releases the 5′LS. To test whether such a rearrangement occurs during 30S subunit assembly, I aimed to reconstitute 30S subunits from the 5′LS•16S complex.

5′LS is removed during 30S subunit reconstitution

After seeing that these RNA-RNA complexes can form, I challenged them with different components, ranging from r-proteins to assembly factors, to test how binding of different factors influenced the stability of 5′LS•16S RNA complex. I formed 5′LS•16S rRNA complexes with labeled 5′LS oligonucleotides at 65℃ for 15 min and then incubated the complexes with TP30 at various concentrations at 42℃ for 15 min. Upon incubation of TP30 with 5′LS•16S rRNA complexes, a super shift can be observed in the native gel after 15 min suggesting that proteins from TP30 can bind to 5′LS•16S RNA complexes (Figure 20A). At higher concentrations of

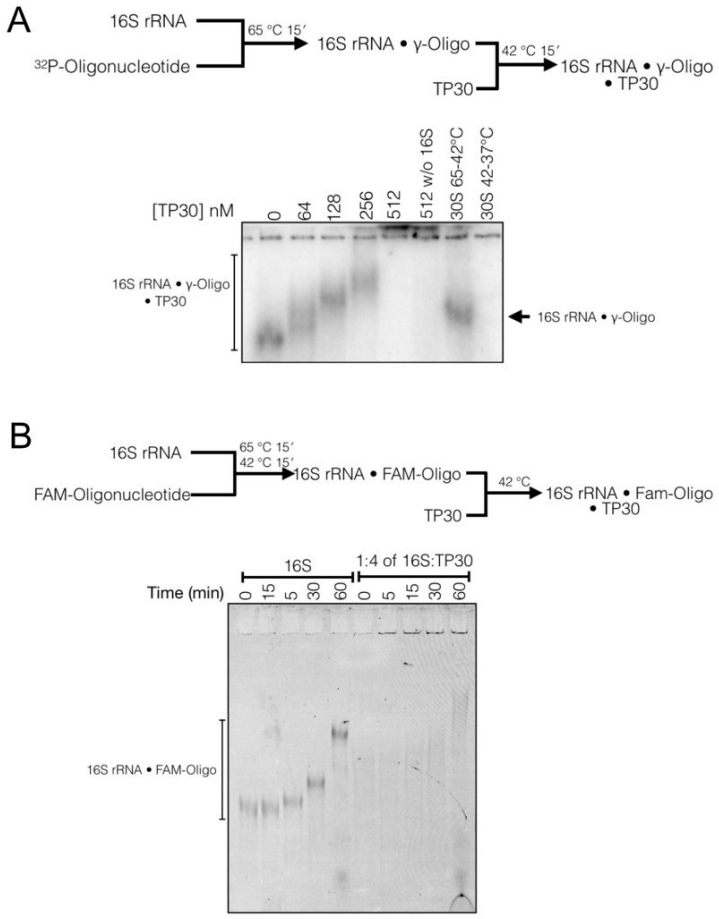


Figure 20: Binding of ribosomal proteins to 5′LS•16S rRNA complex.

A) Native gel resolving the 5′LS•16S complexes containing alt-H1. After annealing the complexes at 65 ℃, they were incubated with increasing concentration of total proteins of 30S (TP30) at 42 ℃. The diagram above the gel highlights the steps performed for this experiment. In the two right most lanes, mature 30S is probed for 5′LS binding at different temperatures.

B) Native gel showing the presence of Alt-H1 complex in the absence (16S) or presence of TP30 (1:4 of 16S:TP30) over time. Using a 1:4 ratio of 16S rRNA to TP30, alt-H1 disappears when compared to no TP30. This result would suggest that having TP30 is good enough to refold alt-H1 to H1. The diagram above the gel highlights the steps performed for this experiment.

TP30 relative to 16S RNA, the complexes migrate slower indicating further protein binding until the bands disappear completely at the highest concentration of TP30 (512 nM; Figure 20A). The absence of signal for the 5′LS•16S RNA complex suggests that alt-H1 has been unwound. 30S subunit reconstitution requires 60 min of incubation at 42 ℃ and I found that the 5′LS remains bound to 16S for the entirety of my time course without TP30. I then performed a time course to test if TP30 requires the full 60 min to remove the 5′LS and found that TP30 disrupts the binding in <5 min (Figure 20B).

A new method for detecting 5′LS rearrangement

Native gel mobility shift assays are great for visualizing interactions in non-denaturing conditions. However, for these experiments, I came across a reproducibility issue, where on some days I would see the removal of the 5′LS from the rRNA while on others I would not. I also saw a difference to the extent 5′LS is removed or stabilized on the 16S rRNA. Therefore, I decided to design a new method for examining 30S subunit reconstitution and 5′LS removal using a pelleting assay system (Figure 21A). In this assay, I first form the 5′LS•16S RNA complex in the same manner as for the native gel assays. Following formation of the 5′LS•16S RNA complex, TP30 was added and incubated for 60 min followed by pelleting the sample mixture over a 40% sucrose cushion, which separates large components from small components. All the unbound proteins and RNA that are not heavy enough are removed in the supernatant, while only the heavy complexes are left behind in the pellet. Following the sucrose cushion, I resolve the samples using SDS-PAGE to detect the presence of proteins and RNAs within the pelleted complexes. Therefore, using this assay I can measure the 5′LS signal disappearance as a function of 30S subunit reconstitution. Instead of running native samples, I am able to resolve the presence of the 5′LS in denaturing conditions which would prevent interference from

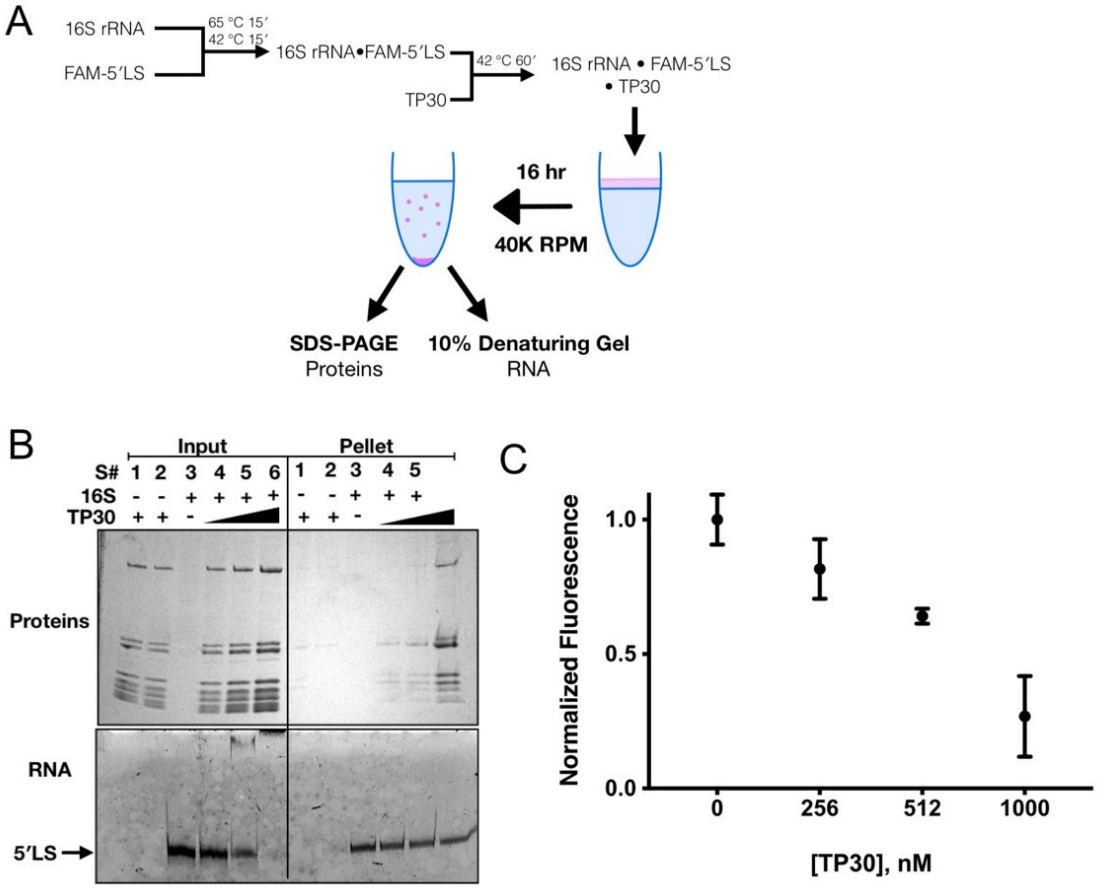


Figure 21: Pelleting assay to detect the reorganization of 5′LS•16S rRNA complex by TP30

A) Schematic outlining the pelleting assay. 16S rRNA and 5′LS are mixed and incubated at two temperatures, then combined with TP30 and incubated for 30 min to allow for reconstitution to occur. Some reconstituted mixture is kept as "input" to be loaded on an SDS-PAGE and a urea denaturing gel, while the rest is loaded on to a 40% sucrose cushion. Large reconstituted complexes pellet, so the sucrose can be discarded, and the pellet is then resuspended and loaded on an SDS-PAGE and 10% denaturing gel.

B) An example of an SDS-PAGE gel resolving the proteins present in the pellet on the top and a 10% denaturing gel to resolve RNA on the bottom. Input sample that is loaded on the sucrose is on the left, while the pellet is on the right. TP30 by itself, sample 1 and 2, does not pellet, but 16S rRNA with 5′LS does. Adding an increasing concentration of TP30 decreases the amount of 5′LS present, as the alt-H1 is being switched to H1.

C) Increasing TP30 concentration decreases the amount of 5′LS that pellets with 16S rRNA. This is based on four separate repeats.

structured rRNAs and proteins. This assay also allows me to examine which proteins are present in the pelleting complexes, which was not possible in the native gel assay. I am also able to compare what was loaded on the sucrose cushion to what pelleted by running the "input" sample and "pellet" sample on the same gel.

Using the pelleting assay, I found that the 5′LS can only pellet when it is bound to 16S rRNA (Figure 22; Lane 1), and TP30 proteins also pellet in the presence of 16S rRNA, showing that 30S subunit reconstitution is occurring (Figure 21B; lane 4-6). TP30 alone cannot pellet over a sucrose cushion because the individual proteins are not heavy enough, resulting in the removal of proteins as can be seen when comparing SDS-PAGE gel between the input and pellet (Figure 21B). Increasing TP30 concentration results in more 30S subunit complexes pelleting, and less 5′LS being bound (Figure 21C). The pelleting assay shows that alt-H1 can be switched to H1 in the presence of only r-proteins and suggests that removal of the 5′LS is likely dependent on some degree of 30S subunit reconstitution. In figure 2, I highlighted three possible proteins in TP30 which might play a role in H1 formation, with S4 and S5 being the most likely candidates. Mutations within r-proteins, such as S5 which binds close to the 5′ of 16S, also resulted in cold sensitivity very similar to the C23U mutants and affects the translation capability of the ribosome (Nashimoto et al. 1971; Roy-Chaudhuri, Kirthi, and Culver 2010). Although S4 is one of the first r-proteins to bind, we know the final bound structure of S4 does not occur until further r-proteins, such as S16, bind later in maturation (Abeysirigunawardena and Woodson 2015). The late stage restructuring of S4 and binding of S5, and their close positions to H1, make them the best candidates for proteins dependent restructuring of H1.

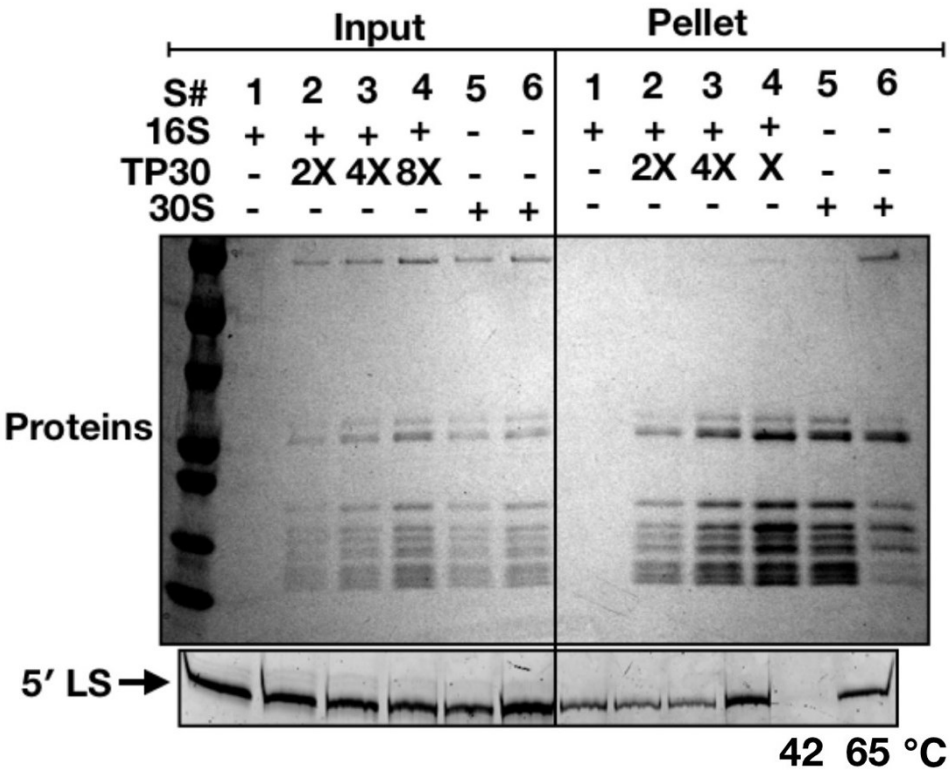


Figure 22: 30S does not bind 5′LS

Sample #5 shows that the 5′LS cannot bind to mature 30S subunits. All the TP30 proteins are present in the 30S sample. However, if 30S rRNA is heated to 65 ℃ and the r-proteins are denatured, 5′LS is able to bind as seen in sample #6. There are some missing r-proteins that did not remain bound to the 16S rRNA after denaturation, showing that the mature 30S structure has been disrupted.

Specific proteins may be required for alt-H1 to H1 transition

As previously seen in native gels, the pelleting assay verified that 5′LS cannot bind to mature 30S subunits (Figure 22; lane 5). This is consistent with the observation that the central pseudoknot present in mature 30S subunit is very stable (Poot et al. 1998). Upon denaturation of 30S subunits at 65°C, I find that 5′LS can now be seen in the pellet, indicating that the 5′LS has bound to 30S subunits (Figure 22; lane 6). These data suggest that alt-H1 cannot be formed unless the complex is denatured under high heat (~ 65°C). My data also shows that denaturing the 30S subunit at 65°C does not fully denature the complex, since some proteins remain bound to 16S rRNA. This suggests that high temperature removes the proteins that are required for keeping H1 folded, and this can help us identify the proteins responsible for alt-H1 to H1 switching.

RbfA's role in alt-H1 to H1 transition

RbfAs ability to rescue the cold sensitive phenotype of 16S rRNA C23U mutation made it a primary candidate for testing its ability to switch alt-H1 to H1. When RbfA is added with TP30 to the 16S rRNA and 5′LS complex, it does not remove more 5′LS than just TP30 alone (Figure 23A; lane 4-6, 9B). This shows that RbfA might not play a direct role in the helix switching that occurs during maturation. The r-proteins are mainly responsible for removing 5′LS and switching alt-H1 to H1. The removal of 5′LS by TP30 is also temperature dependent (Figure 23 C,D). Just like 30S subunit *in vitro* reconstitution requires 42°C, removal of the leader did as well. This highlights that the 16S rRNA and TP30 are reconstituting and removing the 5′LS, switching alt-H1 to H1 as was hypothesized. Keeping the reaction at room temperature did not remove any of the 5′LS.

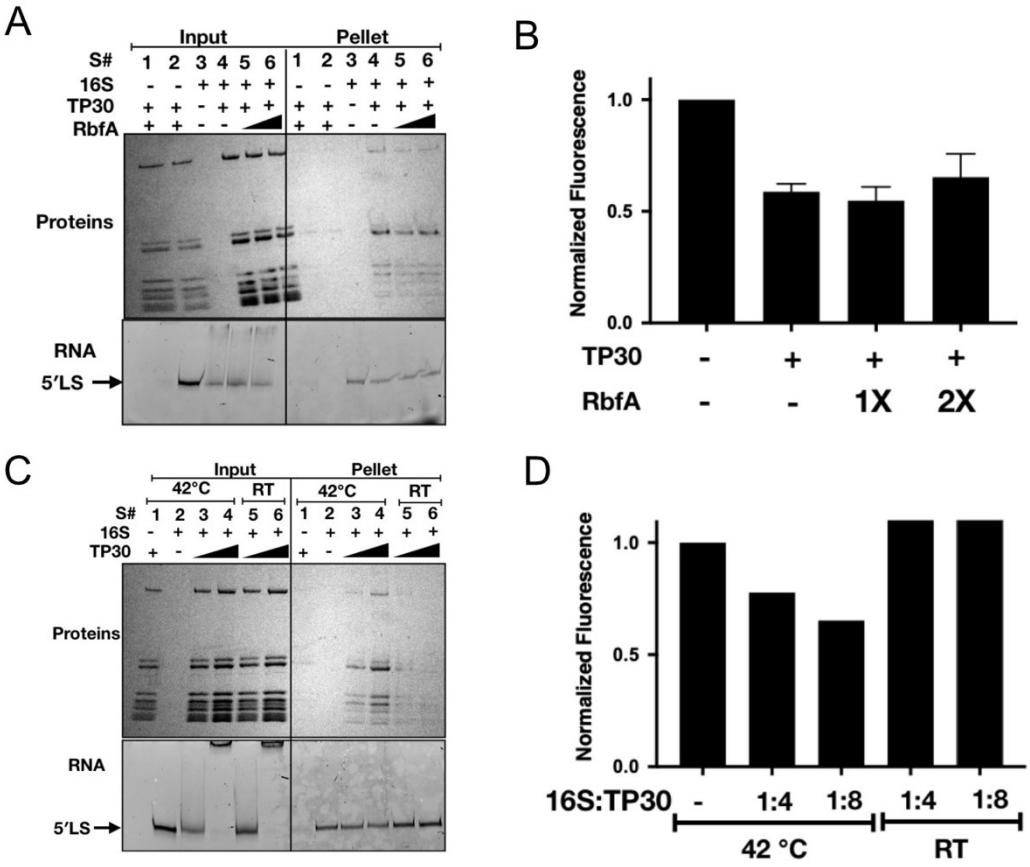


Figure 23: RbfA does not affect the 5′LS•16S rRNA complex and 42 °C is required for removal of 5′LS

A) Gel showing the pelleting of TP30, RbfA, and 5′LS after reconstitution. RbfA, TP30, and 5′LS do not pellet together without the presence of 16S rRNA. Because RbfA is similar in size to other r-proteins, it is difficult to resolve the RbfA band on a 1D SDS-PAGE.
B) TP30 removes 5′LS from 16S rRNA, as seen previously. However, adding 1:1 or 1:2 ratio of 16S:RbfA does not increase the amount of 5′LS removed as was predicted. The amount of 5′LS present with and without RbfA is similar. This is based on three repeats.
C) The removal of 5′LS by TP30 is temperature dependent. Under higher temperature of 42°C, TP30 is removing the 5′LS. At the same temperature, in vitro 30S reconstitution is done, signaling that we are getting reconstituted particles.
D) TP30 at room temperature (RT) does not remove the 5′LS. 30S reconstitution *in vitro* does not occur at RT, therefore alt-H1 is not switched to H1. Reconstitution 30S subunits required 42°C and switching alt-H1 to H1 also requires 42°C. This is based in a single repeat.

Reconstitution is hindered by heat treatment of native 16S rRNA

The pelleting assay was a great way to lower any noise I was detecting on the gel and purify the complexes I was interested in from unbound proteins and RNA. However, I continued to suffer from reproducibility issues, and I tested different temperatures, buffer components for reconstitution, native 16S and TP30 purification and handling, 5′LS variations, but changing all the components in the reaction did not improve the reproducibility issue.

I looked closer at the reconstitution reaction itself that I was performing by doing ultracentrifugation over a 10%-40% sucrose gradient (M. Nomura 1968). I found that the 16S rRNA can reconstitute into mature 30S subunit in the presence of 5′LS at 42℃ (Figure 24). However, if native 16S rRNA is incubated at 65℃ with the 5′LS, it can no longer reconstitute into 30S subunits. This may explain the variation I was seeing. In order to get as much 5′LS bound, I typically incubated the 16S rRNA at 65℃ with the 5′LS. Although 5′LS can bind to 16S rRNA at 42℃, the signal in my gels were weak indicating that some 16S RNA could not bind the 5′LS unless it was denatured. By incubating the RNAs at high temperature, I was denaturing the structure that was present in the purified native 16S rRNA. Once the rRNA was denatured, it was unable to be refolded into a reconstitution-competent conformation. Although other people have shown that *in vitro* transcribed 16S rRNA can reconstitute into 30S particles, the reconstitution is inefficient and other intermediates can be found (Figure 25). By denaturing the 16S rRNA I was getting varied results because I would end up with different populations of 30S particles every experiment due to having so many different components present in my reactions and the varied way in which 30S subunit reconstitutes.

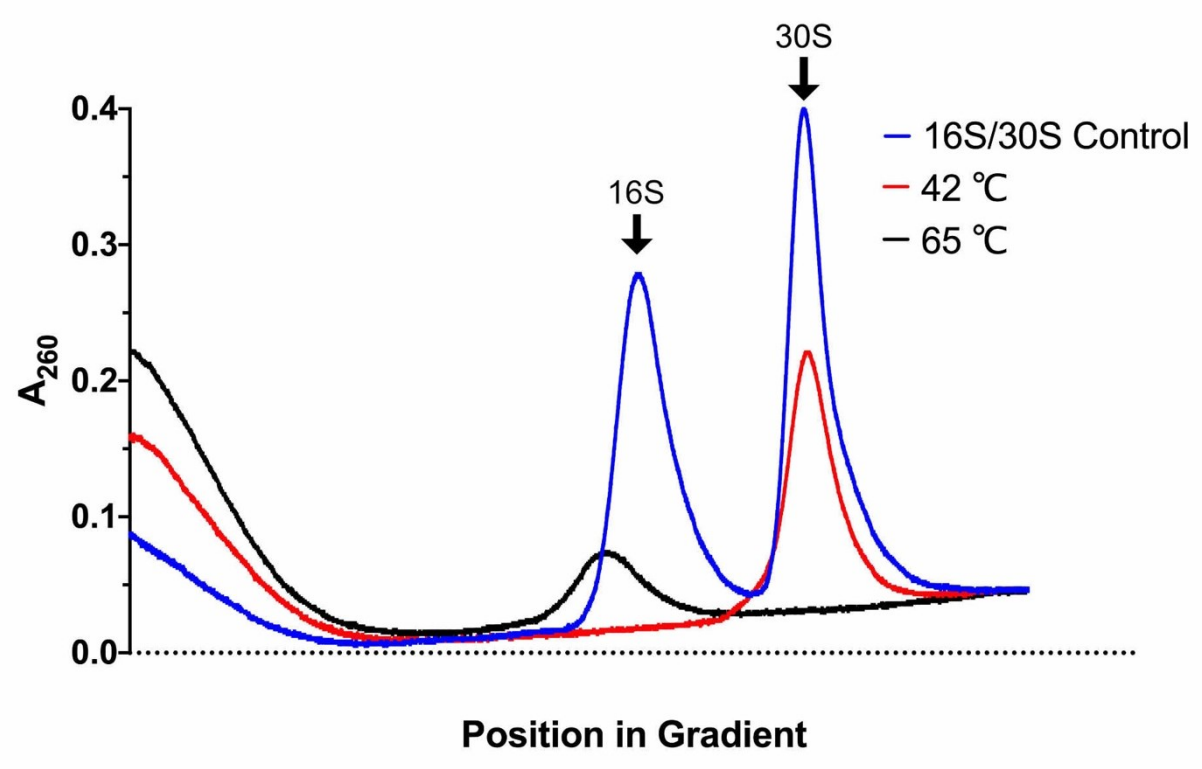


Figure 24: Reconstituting 30S ribosomes after 16S rRNA denaturation in the presence of 5′LS

A 10%-40% sucrose gradient showing the position of 16S rRNA and 30S subunit after reconstitution in the presence of 5′LS. Native 16S rRNA was pre incubated at 42℃ (red) or 65℃ (black) with 5′LS, then TP30 was added at 42℃ and the standard protocol for 30S reconstitution was followed. The peaks were then resolved with a fractionator. The blue line represents a mixture of 16S rRNA and 30S subunit as a control to establish the position of the peaks. After 42℃ incubation and in the presence of the 5LS, 30S particles can be reconstituted. However, by pre-incubating 16S rRNA at 65℃, no reconstitution occurs.

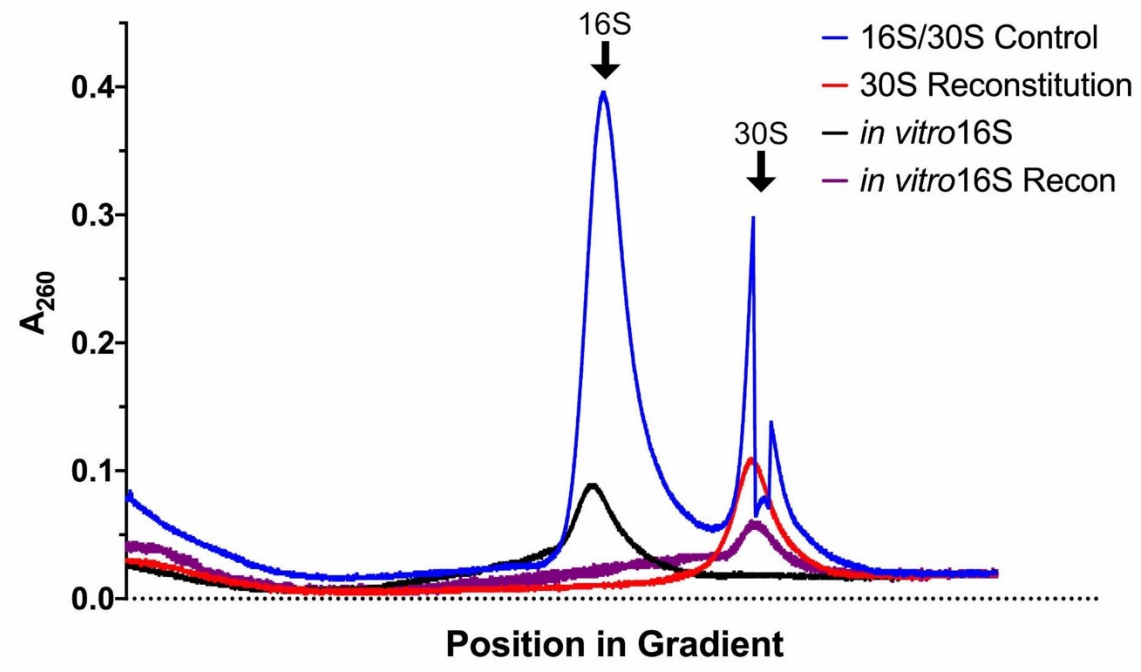


Figure 25: Reconstitution of *in vitro* transcribed 16S

A 10%-40% sucrose gradient showing the position of 16S rRNA and 30S subunit after reconstitution. The blue line represents a mixture of native 16S and 30S that was done as a control to detect the position of the two peaks. The red line is reconstituted 30S from native 16S. The black line is *in vitro* transcribed 16S and it runs at the same position as native 16S. The purple line represents 30S reconstitution with *in vitro* transcribed 16S. Although a large amount of 30S particles do get reconstituted, there is a long tail stretching from the 16S peak which highlights the variation in reconstituting *in vitro* 16S.

Hfq plays a role in 30S subunit biogenesis

Host Factor for replication of bacteriophage Qb (Hfq) is a hexameric RNA chaperone that has a regulatory role in bacteria (Santiago-Frangos and Woodson 2018). Its ability to catalyze the binding of different mRNAs and sRNAs makes it an important regulator of transcription and stability of varied RNAs. Parts of the 5′LS have sequence similarities to sRNAs which bind Hfq (A rich patches), suggesting that Hfq might play a role in the maturation of 30S subunits. To test this hypothesis, I performed binding assays between Hfq and the 5′LS (Figure 26). I found that Hfq binds the 5′LS tightly and their interactions are comparable to those of identified Hfq sRNA substrates (Figure 26A). I then tested mutant Hfq65 which is missing its regulatory C-terminal tails and found that it binds 5′LS as well, but not as tightly as the WT (Figure 26B) (Santiago-Frangos et al. 2016). Hfq65 Y25D which has a mutation on the face of the protein which binds poly-A sequences, is still able to bind the 5′LS, suggesting that this is not the binding site of the 5′LS (Link, Valentin-Hansen, and Brennana 2009). The Hfq65 K31A mutants, which is located on the same face as Y25, bound the 5′LS as well but had less higher order complexes which represent multiple Hfq proteins bound to the 5′LS.

A few months after these experiments were done, a paper was published showing that an Hfq deletion strain has a similar phenotype to an RbfA deletion. The Hfq deletion strain was cold sensitive, had a decrease in 70S ribosomes, and an increase in 30S and 50S subunits. The deletion strain was also prone to mistranslation in the presence of antibiotics, and footprinting results showed a difference in probing at the 5′ helix 1. This suggests that Hfq might play a role in the switching of alt-H1 to H1 during the late stages of maturation, and the stabilization of the central pseudoknot. Although my data suggest that r-proteins are required for 5′LS removal, Hfq may catalyze this process, just as other assembly factors but is not a requirement. It has also been

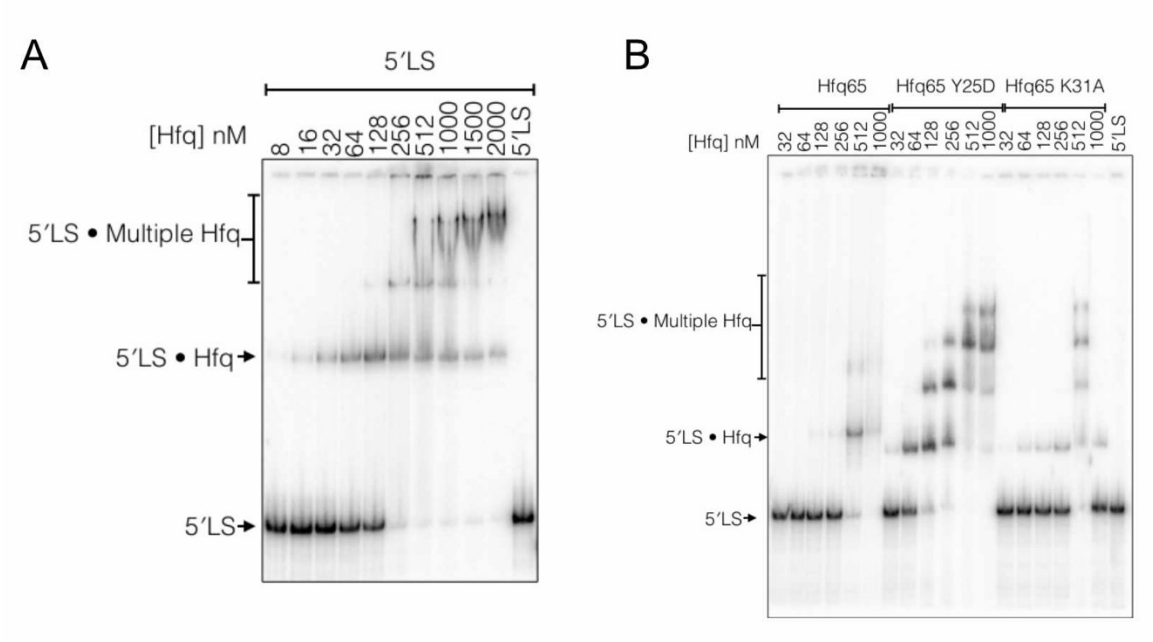


Figure 26: Hfq binds to the 5′LS

A) Native gels showing titration of Hfq proteins to a constant concentration of the 5′LS. Hfq binds the 5′LS efficiently, and the size of bands begin to increase as multiple Hfq proteins bind to the RNA.
B) Native gel comparing the binding of the C-terminus deletion mutant Hfq65, Hfq65 Y25D, and Hfq65 K31A. 5′LS can still bind Hfq, but there are less higher order complexes with multiple Hfqs, as is expected. Y25D does not seem to affect Hfq binding, except increasing the amount of higher order complexes compared to Hfq65. K31A can weakly bind the 5′LS but does not form higher order complexes, just like Hfq65.

found that Hfq stimulates RNase E dependent degradation of RNA, which might hint at the possibility of Hfq recruiting RNase E to cleave the 5′LS (Santiago-Frangos and Woodson 2018; I.M. Sharma, Korman, and Woodson 2018).

Conclusion and future directions

The synthesis and maturation of the ribosome and its ribosomal subunits is a complex process which the cell spends the majority of its energy on (Williamson 2006). Due to its complexity, it is difficult to study and isolate all the different variables that play a role in these processes. I had difficulty in reproducing my experiments because of the complexity involved in 30S subunit folding. Keeping all 21 r-proteins, assembly factors, and two RNAs active proved to be difficult at the same buffer conditions and through an hour of incubation and an overnight ultracentrifugation spin.

Although reproducibility was an issue, there are still some take away points that I have learned. The 5′LS does bind to the 16S rRNA, and it does so in a sequence dependent manner. The same mutation (C-5U) that rescues the cold sensitivity in C23U mutants, also prevents 5′LS from binding to 16S, showing that the structure formed between the two RNAs likely forms in trans *in vitro*. Denaturing native 16S rRNA makes it difficult to reconstitute with TP30, just like it is difficult to reconstitute 30S subunit from *in vitro* transcribed 16S. The purified 16S already has some structure that remains even after the purification step. Although TP30 dependent removal of 5′LS was inconsistent, there have been multiple instances in which TP30 did remove the 5′LS from 16S. In order to get a better idea of what is going on, different TP30 purifications should be analyzed with mass spectrometry in order to see if the current purification method is consistent and gives us the full a stochiometric amount of 21 r-proteins every time. There is a potential for some TP30 preparations to have less S4, S5 and S12, all candidates that might help

H1 fold (Figure 17). The 5′LS cannot bind to mature 30S subunit but can bind to denatured 30S subunit. This shows that the proteins that come off during denaturation are responsible for keeping the central pseudoknot folded and preventing alt-H1 from forming. A mass spectrometry analysis of mature 30S subunit and denatured 30S subunit would eliminate some r-proteins that do not play a role in central pseudoknot stability. Figure 22 shows that there are still proteins bound to denatured 30S subunit and it will be interesting to identify them. When the removal of 5′LS RNA works, it is temperature dependent, and occurs at the same temperature as 30S subunit reconstitution. This tells us that protein and RNA interactions do occur, and a different structure is being stabilized over the competing 5′LS and 16S rRNA structure.

My results suggest that RbfA does not play a role in the altH1/H1 transition, however, further experiments are required to confirm this. Hfq binds to the 5′LS, interacts with RNase E and S12, all components which exist around the central pseudoknot in pre-30S complexes, making Hfq a contender for participating in the alt-H1 to H1 conversion (Figure 17) (Strader et al. 2013; I.M. Sharma, Korman, and Woodson 2018).

These are difficult experiments, but progress can still be made. In our lab, we established and optimized a protocol to purify pre-30S particles. Because these particles contain 17S rRNA, we can purify and probe the structure using footprinting to measure the position of the 5′LS. We could also perform cryo-EM on either the pre-30S particle alone or as part of the 70S complex. We know that pre-30S can go into the translational pool of ribosomes, so it would be interesting to see how the 5′LS is positioned in these complexes (Indra Mani Sharma and Woodson 2019; Soper et al. 2013). We can also try performing reconstitution experiments using natively purified pre-30S complexes. I was unable to get *in vitro* reconstitution of 17S rRNA to work efficiently and knowing that native 16S purifies with a somewhat already folded structure, it

might be easier to reconstitute pre-30S particles from native pre-17S. The reconstitution of native 17S with TP30 purified from pre-30S and 30S subunit can help us understand the differences between the pre-30S and 30S complexes. We can also fluorescently label the 5′LS of native 17S and watch its processing as more reconstitution components are added. The experiments proposed above would help us increase our understanding which proteins are responsible for alt-H1 to H1 switching and would allow us to visualize the processing for pre-30S into mature 30S subunits.

Materials and Methods

30S subunit purification

The 30S subunit purification protocol was adopted from previously published protocols (Nierhaus 1990). MRE600 cells were grown in LB media at 37 ℃, then at OD_{600} =1, kept at 4 ℃ for one hour to break 70S ribosomes into the 30S and 50S subunits. The cells were then pelleted, lysed through emulsification by Emulsiflex Homogenizer and debris cleared. Lysate was then loaded on to a 40% sucrose cushion containing buffer D (1.1 M Sucrose, 20 mM Tris-HCl, pH 7.5, 500 mM NH_4Cl, 10 mM $MgCl_2$ and 0.5 mM EDTA). To purify large complexes (30S, 50S, 70S), the lysate was spun in a Ti 60 rotor, O/N at 40K RPM, 16 hrs, 4 ℃ in the UltraCentrifuge. The next day the supernatant decanted, and the pellet washed with buffer B (20 mM Tris-HCl, pH 7.5, 20 mM NH_4Cl, 0.5 mM $MgCl_2$, 0.5 mM EDTA, and 6 mM β-Me) to get rid of the brown layer on the translucent pellet. The pellet was then resuspended in 500 µl buffer C (50 mM Tris-HCl, pH 7.5, 1.5 mM NH_4Cl, 5 mM $MgCl_2$, and 6 mM β-Me) by shaking on ice for 3 hrs. The resuspended pellet was then dialyzed twice into buffer E (50 mM Tris-HCl, pH 7.5, 150 mM NH_4Cl, 1 mM $MgCl_2$, and 6 mM β-Me) for 30 min each, 60 min total. The solution was then loaded on to a 10%-40% sucrose gradient containing buffer E. To separate 30S subunits from

50S subunits and 70S ribosomes, the mixture was spun in a SW 28 rotor for 16 hrs, at 25K RPM, 4 ℃. The 30S subunits were then separated from the rest using a fractionator, and elution peak detected via A_{260}. In order to concentrate the 30S subunit, it was spun again in Ti60 rotor on a 40% cushion with buffer D at 40K RPM, 4 ℃, 16hrs. The supernatant was decanted and resuspended in Buffer A (20 mM Tris-HCl, pH 7.5, 20 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA, and 6 mM β-Me) over 3 hrs while shaking. The 30S subunit was then aliquoted and frozen.

16S purification

The 16S purification protocol was adopted from previously published protocols (Nierhaus 1990). Buffer containing 20 mM HEPES pH 7.6, 0.5 mM $MgCl_2$, 300 mM NH_4Cl, 1 mM Spermidine, 0.1 mM Spermine, 6 mM β-Me and 1% SDS was combined with 250 µl of 30S subunit and volume brought up to 500 with npH_2O. After mixing, equal volume of phenol was added and vortexed at 15 min in the cold room. The aqueous layer was separated from the organic layer with a 13K RPM spin for 10 minutes. The aqueous layer was separated once again and phenol was added a second time. After this second separation, chloroform was added to the aqueous layer, and again separated in the same way as previously. Then 1.25 µl of 0.5 M EDTA, 50 µl of 3M CH_3COONa and 3x EtOH (~1.7 ml) was added and kept overnight in -80 ℃. The precipitated RNA was spun down at 9K RPM for 30 min, at 4 ℃. A 70% EtOH wash was performed and then the pellet was dried in a speed-vac. The dried pellet was then resuspended in 300 µl TE Buffer and stored at -80 ℃.

TP30 purification

The TP30 purification protocol was adopted from previously published protocols (Nierhaus 1990). Buffer containing 20 mM HEPES pH 7.6, 8 mM $MgCl_2$, 600 mM NH_4Cl, 2

mM Spermidine, 0.2 mM Spermine, 10 mM β-Me, was combined with 250 µl of 30S subunit and volume brought up to 500 with npH_2O. 100 µl of 1M magnesium acetate was added, followed by 2.2 ml of acetic acid. The solution was rocked at 4 ℃ for 1 hr, then centrifuges at 9K RPM for 30 min. 16.5 ml of acetone was added, mixed, and stored in -80 ℃ overnight. The precipitated proteins were then spun at 9K RPM for 30 min, and the pellet resuspended in 1ml of 7M urea, 20 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 150 mM KCl, and 6 mM β-Me. The proteins were then dialyzed in the same buffer overnight. The next day the proteins were dialyzed into a non-urea containing buffer (20 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 150 mM KCl, and 6 mM β-Me) for 90 min. The concentration was measured at A_{230}.

Native gel mobility shift assays

The 5′LS•16S RNA complex was formed by combining 100 nM of 5′LS and 128 nM of 16S rRNA in 30S subunit reconstitution buffer (80 mM K-HEPES (pH 7.6, 300 mM KCl, 20 mM MgCl2, 6 mM β-Me and 0.01% (w/v) Nikkol detergent), with a final volume of 10 µl. The mixture was then incubated at 65 ℃ for 15 min, followed by 42 ℃ for 15 min. To challenge the complex, 256 nM of TP30 was added, mixed, and kept at 42 ℃ for another 1 hr. 10 µl of the mixture is then loaded on a native gel. Resolving the 5′LS•16S RNA complex on the same gel as free 5′LS required making a two layered gel. The bottom layer was a 16% TBE native gel, while the top layer was a 4% TBE gel. The gel was pre-ran for 30 min at 15 W, and then the complexes resolved for two hours at 15 W. Experiments that were done with radio-labelled RNA were then dried and exposed overnight, while those with fluorescently labelled RNA were scanned for the fluorophore on a GE typhoon.

Pelleting assays

As with the native gel mobility assay, The 5′LS•16S RNA complex was formed by combining 100 nM of 5′LS and 128 nM of 16S rRNA in 30S subunit reconstitution buffer (80 mM K-HEPES (pH 7.6, 300 mM KCl, 20 mM MgCl2, 6 mM β-Me and 0.01% (w/v) Nikkol detergent). The mixture was then incubated at 65 ℃ for 15 min, followed by 42 ℃ for 15 min. To challenge the complex, 256 nM of TP30 was added, mixed, and kept at 42 ℃ for another 1 hr. 10 µl of the mixture is then removed as the input, and stored at 4 ℃ overnight, while the other 90 µl were loaded on a 40% sucrose cushion. The mixture was then spun in a SW-50 Rotor at 40K RPM for 16 hours at 4 °C. The pellet was then resuspended in 20 µl of 30S subunit reconstitution buffer. 5 µl of the input and 10 µl of pellet were mixed with equal amount of SDS-Buffer, and then resolved on a 4%-10% SDS-PAGE and then resolved on a 4%-10% SDS-PAGE. While 5 µl of the input and 10 µl of pellet were mixed with equal amount of 2X denaturing dye (7M Urea, 1X TBE, 0.2% Bromophenol Blue, 0.2 % Xylene Cyanol dye) and then resolved on a 10% denaturing gel. The gels were then imaged after staining or scanned for FAM labelled on the GE typhoon.

30S subunit *in vitro* reconstitution

The 30S subunit purification protocol was adopted from previously published protocols (M. Nomura 1968; Masayasu Nomura and Held 1974; Nierhaus 1990; Culver and Noller 1999), and the additions made in our lab. Natively purified 16S rRNA was incubated in 30S subunit reconstitution buffer (80 mM K-HEPES, pH 7.6, 300 mM KCl, 20 mM MgCl2, 6 mM β-Me, and 0.01% (w/v) Nikkol) for 15 min at 42 °C. Then TP30 was added at a ratio of 1:2 of 16S rRNA: TP30 and incubated for 1 hour at 42 °C. The mixture was then loaded on a 10%-40% sucrose gradient buffer (20 mM K-HEPES, pH 7.6, 330 mM KCl, 20 mM MgCl2) and spun in the SW41

rotor, for 15.5 hours, at 32K RPM and 4 °C. The presence of 16S rRNA and 30S subunits were detected using a fractionator and a UV monitor which detected absorbance at A260. The UV monitor was hooked up to a program which detect the signal, which was then exported and plotted in GraphPad Prism 8.

Chapter 4: Ribosomopathies

Introduction

Ribosomes, molecular machines which are responsible for protein synthesis in all living cells, are indispensable for the proper function and development of a cell. Composed of ribosomal RNA (rRNA) and ribosomal proteins (r-proteins), ribosomes synthesize new proteins using messenger RNAs (mRNAs) as blueprints. Proper protein synthesis is essential for cell growth and division. Cells which lack functional ribosomes are unable to perform proper protein synthesis, and therefore undergo growth arrest until the issue is fixed, or undergo cell death. When cells function as one part of a whole, improper protein synthesis can result in improper growth and disease in the whole or the organism. Ribosomopathies are disorders which arises from improper ribosome biogenesis, the pathway for ribosome maturation from rRNA and proteins to a functioning complex in the cell.

Previously it was believed that ribosome biogenesis is an uncompromising aspect of cell growth and development. If a mutation were to occur in any gene that participates in ribosome biogenesis, it was believed to be lethal because proper ribosome biogenesis is crucial for ribosome function, and therefore cell function. However, over the past few decades, a few human diseases have been linked to defective ribosome biogenesis. Many mutations in proteins and RNA that play a role in ribosome biogenesis have been identified as molecular causes for ribosomopathies. Characterizing these disorders as ribosomopathies has been difficult due to the large number of proteins and RNA that participate in ribosome biogenesis and maturation. These same proteins play other roles in the cells metabolism and therefore it is hard to discern if a disease is the result of its role in ribosome biogenesis or other cell functions. Although many mutations have been linked to different disorders, it has been found that one disease can be caused by multiple mutations

in the same or different gene. There have also been no correlations found between mutations and certain phenotypes that present themselves. Taking all of this into account, it becomes obvious that identifying and characterizing ribosomopathies is a difficult task which we have just begun.

Another interesting aspect of ribosomopathies, is the tissue-specific effect of improper ribosome biogenesis. All the cells within patients suffering from ribosomopathies have the same mechanistic dysfunction within their genes. However, different tissues seem to suffer to different extents. Although ribosomes and their function are essential in all cells, some cells seem to be affected more by altered ribosome biogenesis. Many hypotheses have been proposed to understand this phenomenon but a consensus has not been established so far. Some believe that the tissue specific affect is due to different tissue having a different tolerance for the activation of p53 (Farley and Baserga 2016). When ribosome biogenesis is disrupted, the amount of activated p53 in the cell increases, resulting in cell senescence or apoptosis (Farley and Baserga 2016; Bursac et al. 2014; Golomb, Volarevic, and Oren 2014). Another hypothesis is based on the presence of "specialized ribosomes" (Farley and Baserga 2016; Nakhoul et al. 2014). These ribosomes are "specialized" in the sense that they might have modified rRNA and/or have a bound protein that is dependent on a specific tissue (Xue and Barna 2012). These specialized ribosomes will then go on to translate specialized mRNAs. A different hypothesis ponders that this difference in tissue susceptibility is due to tissues having different needs for ribosomes. Fast growing cells and tissues will need more ribosomes and therefore have a shorter timeframe to make them. Therefore, cells which have an altered ribosome biogenesis pathway and are fast growing and reproducing will be more susceptible to the production of improper ribosomes. Due to quick growth, cells will find it hard to overcome bottlenecks in ribosome biogenesis arising from a mutation in the pathway. This can result in the buildup to ribosomal intermediates which can stimulate the nucleolar stress

response, causing either growth arrest or cell death (James et al. 2017). Due to this tissue specificity, ribosomopathies are not currently being treated by focusing on the cause, but instead are contained by focusing on the effects, or the symptoms that arise in patients.

A few prevalent illnesses associated with ribosomopathies are anemias, craniofacial abnormalities, skin discoloration, and a higher chance of cancers (Armistead and Triggs-raine 2014; Nakhoul et al. 2014; Narla and Ebert 2010). Some of these diseases are dominant negative, while others are autosomal recessive. Most of ribosomopathies are rare within the population and are quickly identified in newborns. To understand ribosomopathies better, we first must focus on ribosome biogenesis and identify the areas where the pathway might be susceptible to mutations that disrupt biogenesis, but not necessarily make it impossible.

Eukaryotic Ribosome Biogenesis

Our current understanding of eukaryotic ribosome biogenesis comes from the study of the yeast *Saccharomyces cerevisiae* (Kressler, Hurt, and Ba 2017). Most of the energy consumed by a eukaryotic cell is focused on ribosome biogenesis (James et al. 2017). In Eukaryotes, the ribosome is composed of four rRNA molecules, 18S rRNA in the small subunit (SSU), and 25S, 5.8S, and 5S rRNA in the large subunit (LSU). These rRNAs are bound by 80 r-proteins (Kressler, Hurt, and Ba 2017). The SSU is bound and molded by 33 r-proteins (names starting with RPS), while the LSU is bound by 47 r-proteins (names starting with RPL). These four rRNAs and 80 proteins make up the mature ribosome. However, during ribosome biogenesis, there are other proteins and RNAs that intermittently bind and help to shape, export, and test the ribosome. There are over 200 proteins which bind and come off the ribosome during biogenesis which are called assembly factors (Kressler, Hurt, and Ba 2017; Nerurkar, Altvater, and Gerhardy 2015; Tomecki, Sikorski, and Zakrzewska-placzek 2017). In total, there are more than 350 proteins that are

involved in eukaryotic ribosome biogenesis (Nerurkar, Altvater, and Gerhardy 2015).

The DNA encoding 18S, 25S, and 5.8S rRNAs is transcribed as a single 35S by RNA Polymerase I in the nucleolus. While the DNA encoding the 5S rRNA is located outside the nucleolus and is transcribed by RNA Polymerase III (Kressler, Hurt, and Ba 2017). The r-proteins mRNA is transcribed by RNA Polymerase II, which shows that ribosome biogenesis is dependent on all three RNA Polymerase that are found in eukaryotes. The 35S rRNA forms the backbone for the first ribosome intermediate, the 90S. The 90S intermediate contains most of the r-proteins that will be bound to the SSU or the 40S subunit, while having very little of the LSU r-proteins bound. The 90S is bound by small nucleolar RNA (snoRNA) which form small nucleolar ribonucleoprotein particles (snoRNPs) which perform ribose methylation or pseudouridine formation. Subsequently, the 90S intermediate is processed by nucleases which separate the two components that will form the SSU and the LSU (Tomecki, Sikorski, and Zakrzewska-placzek 2017).

As the ribosome intermediates are formed, they travel from the nucleolus to the nucleoplasm and then get transported to the cytoplasm where they mature into the small and large subunit (Nerurkar, Altvater, and Gerhardy 2015). These two mature subunits will come together to form a functional ribosome. Most of the r-proteins are bound to the SSU intermediate relatively early, and as it travels through the nucleoplasm, a few more r-proteins bind and shape the 20S rRNA intermediate for its final processing to 18S in the cytoplasm. Export proteins begin to bind the pre-40S and allows it to pass through the nuclear pore complex. The pre-40S subunit uses general export proteins with the help of Ran-GTP such as Crm1, Mex67/Mtr2, and Rrp12 (Nerurkar, Altvater, and Gerhardy 2015). No know specific exporter proteins have been identified for the pre-40S (Nerurkar, Altvater, and Gerhardy 2015). These export proteins do not only allow

the pre-40S to enter the cytoplasm but also protect the yet to be structured nucleotides. After export, most assembly factors and export proteins are removed, the pre-40S binds the 60S to form a pre-80S complex. During the formation of this complex, the 20S rRNA is processed to the final 18S rRNA and the 60S is then disassociated (Kressler, Hurt, and Ba 2017). Finally, the whole 40S is checked for proper function by assembly factors which probe for the movement of specific nucleotides, which then goes on to participate in mRNA translation.

Unlike the pre-40S, the pre-60S is not formed at all while in the 90S intermediate. Instead, the pre-60S formation is dependent on being processed from the pre-40S as well as the 3′ end being fully transcribed. LSU r-proteins begin to bind after the rRNA has been fully transcribed with the 5′ and 3′ ends coming together (Kressler, Hurt, and Ba 2017; Tomecki, Sikorski, and Zakrzewska-placzek 2017). As r-protein begin to bind in the nucleolus, the rRNA is processed into the 25S and 5.8S rRNA. The 5S rRNA binds its r-proteins outside the nucleolus and then travels to the nucleolus where it interacts with the pre-60S complex. As this pre-60S subunit moves from the nucleolus to the nucleoplasm, it binds more r-proteins and assembly factors. It also binds general export proteins as mentioned above for the pre-40S, but unlike the pre-40S, a specific export factor Nmd3 with a nuclear export sequence (NES) has been identified for the pre-60S subunit (Nerurkar, Altvater, and Gerhardy 2015). Binding of Nmd3 marks the pre-60S complex for export to the cytoplasm where it is processed to a mature 60S subunit. In the cytoplasm, the pre-60S is bound by additional assembly factors in a specific order to disassociate nucleus bound assembly factors as well as export proteins. While removing bound assembly factors, these new assembly factors test the pre-60S for function before allowing it to enter the pool of mature 60S subunits. Finally, the mature 60S and 40S come together on an mRNA which it can begin to translate as a mature 80S ribosome.

After following the maturation of the eukaryotic ribosome, it becomes clear how mutations in certain proteins or RNAs could have a deleterious effect on the structure and function of the ribosome. Although it is thought that the SSU and LSU can arrive at their mature structure through many paths, only a few of the highly populated intermediate complex have been identified so far, showing us that there are checkpoints that need to be reached with any path chosen. Mutations that might inhibit one pathway or result in checkpoint lag, will produce accumulated ribosome intermediates which will activate the nucleolar stress response. From transcription of ribosomal DNA to protein binding, rRNA processing, nucleus export and assembly factor binding and probing, there are many points at which ribosome biosynthesis can be influenced by unexpected changes in cellular genetics which can lead to ribosomopathies.

Identified Ribosomopathies

There have been approximately ten diseases classified as ribosomopathies whose molecular causes have been identified (Nakhoul et al. 2014; Narla and Ebert 2010). Dimond-Blackfan Anemia is the primary example of a ribosomopathy because it has been recreated in mouse and zebrafish models, and its cause has been shown to only be the result of mutations in r-proteins. All ribosomopathies do not have a long-term cure, instead their symptoms are treated with steroids or bone marrow transplants. People suffering from a ribosomopathy usually have a blood disorder, in which their red blood cells do not form properly, they do not have the proper shape, or lack the progenitor red blood cells. White blood cells have also been affected in some ribosomopathies. Patients also have physical defects as well as potential for mental retardation. Ribosomopathies increase the chances of contracting cancers such as leukemia. The following is a closer look at some of the characterized ribosomopathies.

Dimond-Blackfan Anemia (DBA): DBA arises due to mutations in both large and small subunit

r-proteins (Danilova and Gazda 2015; Nakhoul et al. 2014). Missense, nonsense, splice site mutations, and deletions have been found in genes encoding the *RPS* and *RPL* genes (Danilova and Gazda 2015). All r-proteins mutations that have been characterized in patients are heterozygous. *RPS19* mutations account for 25% of individuals suffering from DBA. These mutations directly affect ribosome biogenesis by influencing r-proteins and their function of shaping the rRNA to which they bind. Mutations in RPS19 and RPS24 have been found to increase the amount of unprocessed 18S rRNA, which result in decrease of mature 40S production (Narla and Ebert 2010). Mutations in RPL genes result in a decrease in 60S production which results in an overall reduction in 80S mature ribosomes.

Patients suffering from DBA are usually diagnosed within the first year of life due to lethargy and a pale complexion. Patients suffer from diseases that arise from a malfunctioning bone marrow such as Macrocytic anemia, hypoplastic anemia and a lack of, or absence of red blood cell precursors (Narla and Ebert 2010). Some patients exhibit physical abnormalities such as short stature, craniofacial and cardiac defects, and thumb abnormalities (Nakhoul et al. 2014). Although 18 r-protein genes have been linked to DBA, no correlation has been made between a mutation and symptoms exhibited by patients (Danilova and Gazda 2015). This suggests that mutations in *RPS19* can lead to different phenotypes. DBA patients are also more susceptible to developing osteosarcoma or bone cancer. The current treatments for DBA is corticosteroids with a bone marrow transplant as a long-term solution.

5q-syndrome: A subtype of myelodysplastic syndrome (MDS) which is characterized by disruption in red blood cell production, 5q-syndrome is the result of deletions in chromosome 5, between bands q21 and q23 (Narla and Ebert 2010). It is one of the few non-congenital ribosomopathies, that arises at random. This deletion results in patients missing one copy of *RPS14*

gene, which is one of the 33 r-protein that bind 18S rRNA in the 40S SSU. Although one copy of *RPS14* is present, the haploinsufficiency is enough to prevent the maturation of the 18S rRNA. Therefore, ribosome biogenesis becomes bottlenecked at the processing of the rRNA, until the cell can translate more Rps14 protein.

Patients suffering from this disease have enlarged red blood cells or macrocytic anemia, and bone marrow cells that do not have the correct nucleus segmentation or hypolobulated megakaryocytes. Patients also have either a normal or higher number of platelets, with a 10% chance of developing leukemia. Blood transfusions is the main treatment that is used to help patients cope, but recently lenalidomide has been shown to decrease the amount of transfusions patients need (Danilova and Gazda 2015). Lenalidomide is an analog of thalidomide that is thought to improve erythropoiesis, erythroid differentiation, and to inhibit p53, preventing the cell from undergoing the nucleolar stress response.

Schwachman-Diamond syndrome (SDS): SDS is the result of a mutation within the *SBDS* gene. In 90% of patients, both copies of the gene are mutated (Nakhoul et al. 2014). SBDS protein is highly conserved in in eukaryotes and is thought to have multiple functions, including mitotic spindle stabilization, stress response, RNA processing as well as ribosome biogenesis. SBDS has been found to be bound to the pre-60S and its function is necessary to progress the maturation of the pre-60S rRNA in the cytoplasm. It is thought to remove EIF6, an export and maturation protein of the pre-60S, from the pre-60S with the help EFL1-GTP (Ruggero and Shimamura 2014). Due to its many predicted functions in the cell, it has been controversial to assign SDS as a ribosomopathy. The phenotypes exhibited by patients might be the result of deregulation of mitotic spindles or a lack of RNA processing instead of ribosome biogenesis. However, lymphoblasts that

we extracted from patients had a defect in ribosome subunit joining, showing that the 60S is not fully matured.

Patients suffer from neutropenia or low count of neutrophils, as well as steatorrhea or excess fat in feces (Narla and Ebert 2010). Their pancreatic exocrine function is also lower than normal, resulting in developmental disorders that lead to short stature, and skeletal deformations with cognitive impairments. Patients also have a higher chance of developing cancer. Current treatments are blood transfusions, pancreatic enzymes, antibiotics with a definitive treatment being bone marrow transplant.

Dyskeratosis Congenital (DC): DC could arise from mutations in nine different genes. The disorder can be autosomal dominant when TERC and TERT are mutated, autosomal recessive when TERT and other regulators of non-coding RNAs are mutated, and X-linked when DKC1 is mutated (Ruggero and Shimamura 2014). Mutations in non-coding RNA regulators makes up 50% of mutated proteins. TERT, which codes for telomerase reverse transcriptase, and TERC, which codes for telomerase RNA, both participate in telomeres maintenance. DKC1 or deskerin, adds specific DNA sequence repeats to the end of chromosomes. Therefore, it is still controversial to include DC in the ribosomopathy category as any identifiable phenotype can be related to the deregulation of telomeres (Narla and Ebert 2010). However, DKS1 is known to associate with snoRNPs which contain *H/ACA* small RNAs, that bind to rRNA during transcription and modify uridines to pseudouridines. Effectively, DKS1 participates in the modification of rRNAs which influence its folding, protein binding, mRNA binding and decoding. The lack of rRNA modification might influence the translation of internal ribosome entry site (IRES) containing mRNAs, such as the p53 mRNA, leading to a decrease in p53.

Patients are usually diagnosed with DC when they contain three symptoms: a rash or skin pigmentation, nail dystrophy, and mucosal leukoplakia or oral white patches. Patients develop multiple cytopenias, where they have a low level of red blood cells, neutrophils, platelets, and granulocytes. Majority of patients die from bone marrow failure and have a larger chance of developing leukemia or solid tumors. Current treatments are steroids and hematopoietic stem cell transplantation. However, due to the strain this disorder puts on other organ systems, stem cell transplantation has a low success rate.

Cartilage Hair Hypoplasia (CHH): This is a rare condition that is found in Old Order Amish population and Finnish decedents. In the general population, this ribosomopathy is uncommon (Nakhoul et al. 2014; Narla and Ebert 2010). The disorder arises due to a mutation in the *RMRP* gene. *RMRP* does not code for a protein, but instead is the RNA part in the RNase mitochondrial RNA-processing complex (RNase-MRP). RMRP is classified as a snoRNA and is localized in the nucleolus. It is thought to operate in the maturation the of 5′ end of the 5.8S rRNA. It is also believed to participate in small interfering RNA (siRNA) synthesis which go on to influence hair development, hematopoietic differentiation and skeletal development (Nakhoul et al. 2014).

Patients who have this disease have a lack of hair, short stature or short limbs, and bone deformities. Patients are more susceptible to the development of cancers, non-Hodgkin lymphoma, and basal cell carcinoma. The current treatment is steroids with stem cell transplant.

Treacher Collins Syndrome (TCS): TCS arises from mutations within *TCOF1* gene. *TCOF1* encodes treacle, a protein that participates in the transcription of rRNAs. It is thought to help RNA polymerase I transcribe ribosomal DNA (Armistead and Triggs-raine 2014; Nakhoul et al. 2014; Narla and Ebert 2010). TCS has also been identified in patients who have mutated RNA polymerase I and III. It is also thought to play a role in the formation of early snRNPs that go on

to modify rRNA during maturation. Patients who have TCS have craniofacial abnormalities due to defects in the formation of the first and second branchial arches. These craniofacial deformities result in troubles breathing, hearing, swallowing and brain development. There are currently no treatments for this disease, but symptoms are addressed through surgeries.

Other Ribosomopathies

There are some ribosomopathies that have not been studied or looked at closer like those described above. Isolated congenital asplenia is the result of a mutation in the *RPSA* gene which codes for an r-protein which binds the SSU (Armistead and Triggs-raine 2014). Aplasia cutis congenital has a mutation in *BMS1*, a ribosomal GTPase. North American Indian childhood cirrhosis (NAIC) is due to a mutation in *CIRH1A* which is a protein that participates in rRNA processing. T cell acute lymphoblastic leukemia arises from mutations in LSU binding r-proteins. Mutations in RPL5, 10, and 22 prevent the proper maturation of the 60S subunit. Patients with Bowen-Conradi syndrome have a mutated *EMG1* gene which plays a role in the methylation of pseudouridines. Anauxetic dysplasia is the result of a mutation in RMRP, just like CHH, which affects pre-rRNA processing.

Role of P53 in Ribosomopathies

The process of making new mature ribosomal subunits is energy and resource intensive for cells and requires constant oversight to make sure the multi-step process is producing proper intermediates and functional end products. The cell uses p53 as a detector for the efficiency of ribosome biogenesis (Golomb, Volarevic, and Oren 2014). If ribosomes are being produced, p53 is tagged for degradation by MDM2. If r-proteins and ribosomal intermediates begin to accumulate, MDM2 is sequestered by unbound r-proteins resulting in p53 stabilization which activates cell cycle arrest and subsequently induce apoptosis. There are current ideas for

ribosomopathy therapy involving the removal of p53 from the cell. However, treating patients by inhibiting p53 activity makes them susceptible to the development of cancers. One of the ways in which it is thought that lenalidomide works in 5q-syndrome patients is by promoting the degradation of p53 (Nakhoul et al. 2014). Mouse and zebra fish models of DBA, 5q-syndrome, and SDS had their erythroid disorder phenotype rescued when p53 was deleted.

In normal functioning cells p53 levels are low and the cell follows its cell cycle. The tumor suppressor p53 is marked for degradation by its E3 ubiquitin ligase MDM2, whose transcription is promoted by p53. It has been found that 5S rRNA in complex with RPL5 and RPL11 sequester MDM2, allowing p53 to become active during the activation of "nucleolar stress response" (Bursac et al. 2014). It has been shown that deleting either RPL5 or RPL11 can stop p53 from becoming activated. Deletion of other r-proteins did not have the same effect on p53, potentially showing that the 5S RNP is essential for detecting ribosome progression. As the concentration of p53 increases in cells, it tetramerizes and binds DNA to induce transcription of p21(and other mRNA) which binds cdk2 to stops the cell cycle. Deletion and mutation in many r-proteins have resulted in p53 activation because it leads to the accumulation of ribosomal intermediates which cannot be further processed and therefore bound by other r-proteins. Active p53 can also inhibit the transcription of rRNA by stopping the activity of RNA polymerase I and III or by activating the retinoblastoma tumor suppressor protein which inhibits rRNA synthesis. p53 can also inhibit ribosome biogenesis by inhibiting mTOR and c-MYC, and by inhibiting the transcription of ribosome export protein CRM1.

Concluding Remarks

Ribosome biogenesis is a major part of the cell life cycle and its efficient performance is required for cell growth and division. Previously, it was thought that ribosome biogenesis is an

uncompromising aspect of life because protein synthesis is required for life to thrive. The discovery of ribosomopathies has highlighted the efficiency of ribosome biogenesis and just how important it is for a cell and an organism to have properly formed and matured ribosomes. Although an organism can survive with alterations in ribosome biogenesis, it would not be able to prosper. There are currently very little treatments available for patients, and potential future development in CRISPR can allow the possibility of a permanent genetic cure for the disease.

References

Abeysirigunawardena, Sanjaya C, and Sarah A Woodson. 2015. “Differential Effects of Ribosomal Proteins and Mg 2 + Ions on a Conformational Switch during 30S Ribosome,” 1859–65. https://doi.org/10.1261/rna.051292.115.1.

Adilakshmi, Tadepalli, Deepti L Bellur, and Sarah A Woodson. 2008. “Concurrent Nucleation of 16S Folding and Induced Fit in 30S Ribosome Assembly” 455 (October): 18–22. https://doi.org/10.1038/nature07298.

Adilakshmi, Tadepalli, Priya Ramaswamy, and Sarah A Woodson. 2005. “Protein-Independent Folding Pathway of the 16 S RRNA 5′Domain,” 508–19. https://doi.org/10.1016/j.jmb.2005.06.020.

Apirion, David, and Peter Gegenheimer. 1981. “Processing of Bacterial RNA.” *FEBS Letters* 125 (1): 1–9. https://doi.org/10.1016/0014-5793(81)80984-2.

Armistead, Joy, and Barbara Triggs-raine. 2014. “Diverse Diseases from a Ubiquitous Process : The Ribosomopathy Paradox.” *FEBS Letters* 588 (9): 1491–1500. https://doi.org/10.1016/j.febslet.2014.03.024.

Balzer, Monika, and Rolf Wagner. 1998. “Mutations in the Leader Region of Ribosomal RNA Operons Cause Structurally Defective 30 S Ribosomes as Revealed by in Vivo Structural Probing.”

Besançon, Willi, Rolf Wagner, Physikalische Biologie, Heinrich-heine-universität Düsseldorf, and D- Düsseldorf. 1999. “Characterization of Transient RNA – RNA Interactions Important for the Facilitated Structure Formation of Bacterial Ribosomal 16S RNA” 27 (22): 4353–62.

Brink, Marcel F, Martin Ph Verbeet, and Herman A De Boer. 1993. “Formation of the Central

Pseudoknot in 1 6S RRNA Is Essential for Initiation of Translation" 12 (10): 3987–96.

Brodersen, Ditlev E., William M. Clemons, Andrew P. Carter, Brian T. Wimberly, and V. Ramakrishnan. 2002. "Crystal Structure of the 30 S Ribosomal Subunit from Thermus Thermophilus: Structure of the Proteins and Their Interactions with 16 S RNA." *Journal of Molecular Biology* 316 (3): 725–68. https://doi.org/10.1006/jmbi.2001.5359.

Brosius, Jcjrgen, Axel Ullrich, Thomas J Dull, and R Robin. 1981. "Construction and Fine Mapping of Recombinant the RrnB Ribosomal RNA Operon Plasmids of E . Co / I" 118: 112–18.

Buck, Janina, Boris Fürtig, Jonas Noeske, Jens Wönert, and Harald Schwalbe. 2007. "Time-Resolved NMR Methods Resolving Ligand-Induced RNA Folding at Atomic Resolution." *Proceedings of the National Academy of Sciences of the United States of America* 104 (40): 15699–704. https://doi.org/10.1073/pnas.0703182104.

Bunner, Anne E, Stefan Nord, P Mikael Wikström, and James R Williamson. 2010. "The Effect of Ribosome Assembly Cofactors on In Vitro 30S Subunit Reconstitution." *Journal of Molecular Biology* 398 (1): 1–7. https://doi.org/10.1016/j.jmb.2010.02.036.

Bursac, Sladana, Maja Cokaric, Giulio Donati, and Sinisa Volarevic. 2014. "Biochimica et Biophysica Acta Activation of the Tumor Suppressor P53 upon Impairment of Ribosome Biogenesis ☆." *BBA - Molecular Basis of Disease* 1842 (6): 817–30. https://doi.org/10.1016/j.bbadis.2013.08.014.

Carter, Andrew P, William M Clemons, Ditlev E Brodersen, Robert J Morgan-warren, Brian T Wimberly, and V Ramakrishnan. 2000. "Functional Insights from the Structure of the 30S Ribosomal Subunit and Its Interactions with Antibiotics." *Nature* 407 (September): 340–48.

Condon, N, Justina Voulgaris, Dmitry Zaporojets, Binghua Shen, Michaal Al-omar, Craig

Squires, and Catherine L Squires. 1999. “Construction and Initial Characterization of Escherichia Coli Strains with Few or No Intact Chromosomal RRNA Operons” 181 (12): 3803–9.

Culver, Gloria M., and Harry F. Noller. 1999. “Efficient Reconstitution of Functional Escherichia Coli 30S Ribosomal Subunits from a Complete Set of Recombinant Small Subunit Ribosomal Proteins.” *Rna* 5 (6): 832–43. https://doi.org/10.1017/S1355838299990714.

Dahlberg, James E. 1999. “An Escherichia Coli Strain with All Chromosomal RRNA Operons Inactivated : Complete Exchange of RRNA Genes between Bacteria” 96 (March): 1971–76.

Dammel, Carol S., and Harry F. Noller. 1995. “Suppression of a Cold-Sensitive Mutation in 16S RRNA by Overexpression of a Novel Ribosome-Binding Factor , RbfA.” *Genes & Development*, 626–37.

Dammel, Carol S., and Harry F Noller. 1993. “A Cold-Sensitive Mutation in 16S RRNA Provides Evidence for Helical Switching in Ribosome Assembly.” *Genes & Development*, 660–70.

Danilova, Nadia, and Hanna T Gazda. 2015. “Ribosomopathies : How a Common Root Can Cause a Tree of Pathologies,” 1013–26. https://doi.org/10.1242/dmm.020529.

Dodd, J, J M Kolb, and M Nomura. 1991. “Lack of Complete Cooperativity of Ribosome Assembly in Vitro and Its Possible Relevance to in Vivo Ribosome Assembly and the Regulation of Ribosomal Gene Expression,” 757–67.

Draper, David E., Dan Grilley, and Ana Maria Soto. 2005. “Ions and RNA Folding.” *Annual Review of Biophysics and Biomolecular Structure* 34: 221–43. https://doi.org/10.1146/annurev.biophys.34.040204.144511.

Dunkle, Jack A., Liqun Xiong, Alexander S. Mankin, and Jamie H D Cate. 2010. “Structures of the Escherichia Coli Ribosome with Antibiotics Bound near the Peptidyl Transferase Center Explain Spectra of Drug Action.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (40): 17152–57. https://doi.org/10.1073/pnas.1007988107.

Dupuis, Nicholas F., Erik D. Holmstrom, and David J. Nesbitt. 2014. “Molecular-Crowding Effects on Single-Molecule RNA Folding/Unfolding Thermodynamics and Kinetics.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (23): 8464–69. https://doi.org/10.1073/pnas.1316039111.

Farley, Katherine I, and Susan J Baserga. 2016. “RNA UK 2016 RNA UK 2016 Probing the Mechanisms Underlying Human Diseases in Making Ribosomes,” 1035–44. https://doi.org/10.1042/BST20160064.

Gaines, Colin S., Timothy J. Giese, and Darrin M. York. 2019. “Cleaning Up Mechanistic Debris Generated by Twister Ribozymes Using Computational RNA Enzymology.” Research-article. *ACS Catalysis* 9 (7): 5803–15. https://doi.org/10.1021/acscatal.9b01155.

Ganser, Laura R., Megan L. Kelly, Daniel Herschlag, and Hashim M. Al-Hashimi. 2019. “The Roles of Structural Dynamics in the Cellular Functions of RNAs.” *Nature Reviews Molecular Cell Biology* 20 (August): 25–27. https://doi.org/10.1038/s41580-019-0136-0.

Gebetsberger, Jennifer, and Ronald Micura. 2017. “Unwinding the Twister Ribozyme: From Structure to Mechanism.” *Wiley Interdisciplinary Reviews: RNA* 8 (3): 1–14. https://doi.org/10.1002/wrna.1402.

Golomb, Lior, Sinisa Volarevic, and Moshe Oren. 2014. “P53 and Ribosome Biogenesis Stress : The Essentials” 588: 2571–79. https://doi.org/10.1016/j.febslet.2014.04.014.

Goto, Simon, Shingo Kato, Takatsugu Kimura, Akira Muto, and Hyouta Himeno. 2011. “RsgA

Releases RbfA from 30S Ribosome during a Late Stage of Ribosome Biosynthesis" 30 (1): 104–14. https://doi.org/10.1038/emboj.2010.291.

Gourse, Richard L, Herman A De Boer, and Masayasu Nomura S. 1986. "DNA Determinants of RRNA Synthesis in E . Coli : Growth Rate Dependent Regulation , Feedback Inhibition , Upstream Activation , Antitermination" 44: 197–205.

Gupta, Neha, and Gloria M Culver. 2014. "Multiple in Vivo Pathways for Escherichia Coli Small Ribosomal Subunit Assembly Occur on One Pre-RRNA." *Nature Publishing Group* 21 (10): 937–43. https://doi.org/10.1038/nsmb.2887.

Held, William A., Byron Ballou, Shoji Mizushima, and Masayasu Nomura. 1974. "Assembly Proteins Mapping of 30 S Ribosomal from Escherichia Coli" 249 (10): 3103–11.

Hellenkamp, Björn, Sonja Schmid, Olga Doroshenko, Oleg Opanasyuk, Ralf Kühnemuth, Soheila Rezaei Adariani, Benjamin Ambrose, et al. 2018. "Precision and Accuracy of Single-Molecule FRET Measurements—a Multi-Laboratory Benchmark Study." *Nature Methods* 15 (9): 669–76. https://doi.org/10.1038/s41592-018-0085-0.

Higgs, Paul G., and Niles Lehman. 2015. "The RNA World: Molecular Cooperation at the Origins of Life." *Nature Reviews Genetics* 16 (1): 7–17. https://doi.org/10.1038/nrg3841.

Hua, Boyang, Kyu Young Han, Ruobo Zhou, Hajin Kim, Xinghua Shi, Sanjaya C Abeysirigunawardena, Ankur Jain, et al. 2014. "An Improved Surface Passivation Method for Single-Molecule Studies." *Nature Methods* 11 (12): 1233–36. https://doi.org/10.1038/nmeth.3143.

Hua, Boyang, Subrata Panja, Yanbo Wang, Sarah A. Woodson, and Taekjip Ha. 2018. "Mimicking Co-Transcriptional RNA Folding Using a Superhelicase." Rapid-communication. *Journal of the American Chemical Society* 140 (32): 10067–70.

https://doi.org/10.1021/jacs.8b03784.

Hulscher, Ryan M, Jen Bohon, Mollie C Rappé, Sayan Gupta, Rhijuta D Mello, Michael Sullivan, Corie Y Ralston, Mark R Chance, and Sarah A Woodson. 2016. "Probing the Structure of Ribosome Assembly Intermediates in Vivo Using DMS and Hydroxyl Radical Footprinting." *Methods* 103: 49–56. https://doi.org/10.1016/j.ymeth.2016.03.012.

Inoue, Koichi, Jingqiu Chen, Qian Tan, and Masayori Inouye. 2006. "Era and RbfA Have Overlapping Function in Ribosome Biogenesis in Escherichia Coli." *Journal of Molecular Microbiology and Biotechnology* 11 (1–2): 41–52. https://doi.org/10.1159/000092818.

Jacob, Asha Ivy, Caroline Köhrer, Bryan William Davies, Uttam Lal RajBhandary, and Graham Charles Walker. 2013. "Conserved Bacterial RNase YbeY Plays Key Roles in 70S Ribosome Quality Control and 16S RRNA Maturation." *Molecular Cell* 49 (3): 427–38. https://doi.org/10.1016/j.molcel.2012.11.025.

James, Allison, Yubo Wang, Himanshu Raje, Raphyel Rosby, Patrick Dimario, Allison James, Yubo Wang, et al. 2017. "Nucleolar Stress with and without P53 Nucleolar Stress with and without P53" 1034 (June). https://doi.org/10.4161/nucl.32235.

Jimenez, Randi M., Julio A. Polanco, and Andrej Lupták. 2015. "Chemistry and Biology of Self-Cleaving Ribozymes." *Trends in Biochemical Sciences* 40 (11): 648–61. https://doi.org/10.1016/j.tibs.2015.09.001.

Jones, Pamela G., and Masayori Inouye. 1996. "RbfA, a 30S Ribosomal Binding Factor, Is a Cold-Shock Protein Whose Absence Triggers the Cold-Shock Response." *Molecular Microbiology* 21 (6): 1207–18. https://doi.org/10.1111/j.1365-2958.1996.tb02582.x.

Kilburn, Duncan, Joon Ho Roh, Liang Guo, Robert M. Briber, and Sarah A. Woodson. 2010. "Molecular Crowding Stabilizes Folded RNA Structure by the Excluded Volume Effect."

Journal of the American Chemical Society 132 (25): 8690–96. https://doi.org/10.1021/ja101500g.

Kim, Hajin, Sanjaya C Abeysirigunawarden, Ke Chen, Megan Mayerle, and Kaushik Ragunathan. 2014. "Protein-Guided RNA Dynamics during Early Ribosome Assembly." https://doi.org/10.1038/nature13039.

Klán, Petr, Tomáš Šolomek, Christian G. Bochet, Aurélien Blanc, Richard Givens, Marina Rubina, Vladimir Popik, Alexey Kostikov, and Jakob Wirz. 2013. "Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy." *Chemical Reviews* 113 (1): 119–91. https://doi.org/10.1021/cr300177k.

Kressler, Dieter, Ed Hurt, and Jochen Ba. 2017. "Review A Puzzle of Life : Crafting Ribosomal Subunits" xx: 1–15. https://doi.org/10.1016/j.tibs.2017.05.005.

Krzyzosiak, W., C. W. Gehrke, R. Denman, K. Nurse, W. Hellmann, M. Boublik, J. Ofengand, and P. F. Agris. 1987. "In Vitro Synthesis of 16S Ribosomal RNA Containing Single Base Changes and Assembly into a Functional 30S Ribosome." *Biochemistry* 26 (8): 2353–64. https://doi.org/10.1021/bi00382a042.

Li, Suzanne C., Catherine L. Squires, and Craig Squires. 1984. "Antitermination of E. Coli RRNA Transcription Is Caused by a Control Region Segment Containing Lambda Nut-like Sequences." *Cell* 38 (3): 851–60. https://doi.org/10.1016/0092-8674(84)90280-0.

Li, Zhongwei, Shilpa Pandit, and Murray P. Deutscher. 1999. "RNase G (CafA Protein) and RNase E Are Both Required for the 5' Maturation of 16S Ribosomal RNA." *EMBO Journal* 18 (10): 2878–85. https://doi.org/10.1093/emboj/18.10.2878.

Link, Todd M., Poul Valentin-Hansen, and Richard G. Brennana. 2009. "Structure of Escherichia Coli Hfq Bound to Polyriboadenylate RNA." *Proceedings of the National*

Academy of Sciences of the United States of America 106 (46): 19292–97. https://doi.org/10.1073/pnas.0908744106.

Liu, Yijin, Timothy J. Wilson, Scott A. McPhee, and David M J Lilley. 2014. "Crystal Structure and Mechanistic Investigation of the Twister Ribozyme." *Nature Chemical Biology* 10 (9): 739–44. https://doi.org/10.1038/nchembio.1587.

McKinney, Sean A., Chirlmin Joo, and Taekjip Ha. 2006. "Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling." *Biophysical Journal* 91 (5): 1941–51. https://doi.org/10.1529/biophysj.106.082487.

Melnikov, Sergey, Adam Ben-Shem, Nicolas Garreau De Loubresse, Lasse Jenner, Gulnara Yusupova, and Marat Yusupov. 2012. "One Core, Two Shells: Bacterial and Eukaryotic Ribosomes." *Nature Structural and Molecular Biology* 19 (6): 560–67. https://doi.org/10.1038/nsmb.2313.

Mizushima, Shoji, and Masayasu Nomura. 1970. "Assembly Mapping of 30S Ribosomal Proteins from E. Coli.Pdf." *Naturet*, 1214–18.

Moore, Peter B., and Thomas A. Steitz. 2002. "The Structural Basis of Large Ribosomal Subunit Function." *Annual Review of Biochemistry*, 813–50. https://doi.org/10.1146/annurev.biochem.71.110601.135450.

Nakhoul, Hani, Jiangwei Ke, Xiang Zhou, Wenjuan Liao, Shelya X Zeng, and Hua Lu. 2014. "Clinical Medicine Insights : Blood Disorders Ribosomopathies : Mechanisms of Disease," 7–16. https://doi.org/10.4137/CMBD.S16952.RECEIVED.

Narla, Anupama, and Benjamin L Ebert. 2010. "Review Article Ribosomopathies : Human Disorders of Ribosome Dysfunction" 115 (16): 3196–3206. https://doi.org/10.1182/blood-2009-10-178129.

Nashimoto, H., W. Held, E. Kaltschmidt, and M. Nomura. 1971. “Structure and Function of Bacterial Ribosomes. XII. Accumulation of 21 s Particles by Some Cold-Sensitive Mutants of Escherichia Coli.” *Journal of Molecular Biology* 62 (1): 121–38. https://doi.org/10.1016/0022-2836(71)90135-5.

Nerurkar, Purnima, Martin Altvater, and Stefan Gerhardy. 2015. “Eukaryotic Ribosome Assembly and Nuclear Export” 319: 107–40. https://doi.org/10.1016/bs.ircmb.2015.07.002.

Nierhaus, Knud H. 1990. *Reconstitution of Ribosomes*. IRL Press at Oxford University Press.

Nikolaev, Nikolai, Lorenzo Silengo, and David Schlessinger. 1973. “Ribosomal Ribonucleic Acid and Messenger Ribonucleic Acid Precursors in Escherichia Coli.” *JBC* 248 (3): 977–84.

Nissen, Poul, Jeffrey Hansen, Nenad Ban, Peter B Moore, and Thomas A Steitz. 2000. “The Structural Basis of Ribosome Activity in Peptide Bond Synthesis” 289 (August): 920–31.

Noller, Harry F. 1991. “Ribosomal RNA and Translation.” *Annu. Rev. Biochem.*, 191–227.

Noller, Harry F., Vernita Hoffarth, and Ludwika Zimniak. 1992. “Unusual Resistance of Peptidyl Transferase to Protein Extraction Procedures.” *Science* 256 (5062): 1416–19. https://doi.org/10.1126/science.1604315.

Noller, Harry F. 2017. “The Parable of the Caveman and the Ferrari : Protein Synthesis and the RNA World.”

Nomura, M. 1968. “RECONSTITUTION OF FUNCTIONALLY ACTIVE 30S RIBOSOMAL PARTICLES FROM RNA AND PROTEINS * RNA and a Number of Different Protein Molecules . The Complexity of These,” 777–84.

Nomura, Masayasu, and William A Held. 1974. *Reconstitution of Ribosomes : Studies of Ribosome Structure , Function and Assembly*.

Panja, Subrata, Boyang Hua, Diego Zegarra, Taekjip Ha, and Sarah A. Woodson. 2017a. “Metals Induce Transient Folding and Activation of the Twister Ribozyme.” *Nature Chemical Biology* 13 (10): 1109–14. https://doi.org/10.1038/nchembio.2459.

———. 2017b. “Metals Induce Transient Folding and Activation of the Twister Ribozyme.” *Nature Chemical Biology* 13 (10): 1109–14. https://doi.org/10.1038/nchembio.2459.

Panja, Subrata, Rakesh Paul, Marc M. Greenberg, and Sarah A. Woodson. 2015. “Light-Triggered RNA Annealing by an RNA Chaperone.” *Angewandte Chemie - International Edition* 54 (25): 7281–84. https://doi.org/10.1002/anie.201501658.

Pedro Lopez-Alonso, Jorge, Tatsuya Kaminishi, Takeshi Kikuchi, Yuya Hirata, Idoia Iturrioz, Neha Dhimole, Andreas Schedlbauer, et al. 2017. “RsgA Couples the Maturation State of the 30S Ribosomal Decoding Center to Activation of Its GTPase Pocket.” *Nucleic Acids Research* 45 (11): 6945–59. https://doi.org/10.1093/nar/gkx324.

Poot, Raymond A, Sjoerd H E Van Den Worm, Cornelis W A Pleij, and Jan Van Duin. 1998. “Base Complementarity in Helix 2 of the Central Pseudoknot in 16S RRNA Is Essential for Ribosome Functioning” 26 (2): 549–53.

Powers, Ted, Gary Daubresse, and Harry F. Noller. 1993. “Dynamics of In Vitro Assembly of 16S RRNA into 30S Ribosomal Subunits.” *Journal of Molecular Biology* 232: 362–74. https://doi.org/140362.

Ramaswamy, Priya, and Sarah A. Woodson. 2009. “S16 Throws a Conformational Switch during Assembly of 30S 5′ Domain.” *Nature Structural and Molecular Biology* 16 (4): 438–45. https://doi.org/10.1038/nsmb.1585.

Razi, Aida, Alba Guarné, and Joaquin Ortega. 2017. “The Cryo-EM Structure of YjeQ Bound to the 30S Subunit Suggests a Fidelity Checkpoint Function for This Protein in Ribosome

Assembly." *Proceedings of the National Academy of Sciences of the United States of America* 114 (17): E3396–3403. https://doi.org/10.1073/pnas.1618016114.

Razi, Aida, and Joaquin Ortega. 2017. "Ribosomal Proteins : Their Role in the Assembly , Structure and Function of the Ribosome," 1–12. https://doi.org/10.1002/9780470015902.a0000535.pub2.

Ren, Aiming, Marija Košutić, Kanagalaghatta R. Rajashankar, Marina Frener, Tobias Santner, Eric Westhof, Ronald Micura, and Dinshaw J. Patel. 2014. "In-Line Alignment and Mg2+ Coordination at the Cleavage Site of the Env22 Twister Ribozyme." *Nature Communications* 5: 1–10. https://doi.org/10.1038/ncomms6534.

Rodgers, Margaret L., Allison L. Didychuk, Samuel E. Butcher, David A. Brow, and Aaron A. Hoskins. 2016. "A Multi-Step Model for Facilitated Unwinding of the Yeast U4/U6 RNA Duplex." *Nucleic Acids Research* 44 (22): 10912–28. https://doi.org/10.1093/nar/gkw686.

Romani, Andrea M P. 2011. "Cellular Magnesium Homeostasis." *Archives of Biochemistry and Biophysics* 512 (1): 1–23. https://doi.org/10.1016/j.abb.2011.05.010.

Roth, Adam, Zasha Weinberg, Andy G.Y. Chen, Peter B. Kim, Tyler D. Ames, and Ronald R. Breaker. 2014. "A Widespread Self-Cleaving Ribozyme Class Is Revealed by Bioinformatics." *Nature Chemical Biology* 10 (1): 56–60. https://doi.org/10.1038/nchembio.1386.

Roy-Chaudhuri, Biswajoy, Narayanaswamy Kirthi, and Gloria M Culver. 2010. "Appropriate Maturation and Folding of 16S RRNA during 30S Subunit Biogenesis Are Critical for Translational Fidelity" 107 (10). https://doi.org/10.1073/pnas.0912305107.

Roy, Rahul, Sungchul Hohng, and Taekjip Ha. 2008. "A Practical Guide to Single-Molecule FRET." *Nature Methods* 5 (6): 507–16. https://doi.org/10.1038/nmeth.1208.

Ruggero, Davide, and Akiko Shimamura. 2014. “Marrow Failure : A Window into Ribosome Biology” 124 (18): 2784–93. https://doi.org/10.1182/blood-2014-04-526301.2784.

Santiago-Frangos, Andrew, Kumari Kavita, Daniel J. Schu, Susan Gottesman, and Sarah A. Woodson. 2016. “C-Terminal Domain of the RNA Chaperone Hfq Drives SRNA Competition and Release of Target RNA.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (41): E6089–96. https://doi.org/10.1073/pnas.1613053113.

Santiago-Frangos, Andrew, and Sarah A. Woodson. 2018. “Hfq Chaperone Brings Speed Dating to Bacterial SRNA.” *Wiley Interdisciplinary Reviews: RNA* 9 (4): 1–16. https://doi.org/10.1002/wrna.1475.

Sarmientos, Paolo, James E Sylvester, Sara Contente, and Michael Cashel. 1983. “Stringent Control of the Tandem E . Coli Ribosomal RNA Promoters from the RrnA Operon Expressed In Vivo in Multicopy Plasmids the Promoters” 32 (April).

Schäferkordt, Jan, Rolf Wagner, Physikalische Biologie, Heinrich-heine-universität Düsseldorf, and D- Düsseldorf. 2001. “Effects of Base Change Mutations within an Escherichia Coli Ribosomal RNA Leader Region on RRNA Maturation and Ribosome Formation” 29 (16): 3394–3403.

Schroeder, Renée, Andrea Barta, and Katharina Semrad. 2004. “Strategies for RNA Folding and Assembly.” *Nature Reviews Molecular Cell Biology* 5 (11): 908–19. https://doi.org/10.1038/nrm1497.

Schroeder, Renée, Rupert Grossberger, Andrea Pichler, and Christina Waldsich. 2002. “RNA Folding in Vivo.” *Current Opinion in Structural Biology* 12 (3): 296–300. https://doi.org/10.1016/S0959-440X(02)00325-1.

Shajani, Zahra, Michael T Sykes, and James R Williamson. 2011. “Assembly of Bacterial Ribosomes.” https://doi.org/10.1146/annurev-biochem-062608-160432.

Sharma, I.M., A. Korman, and S.A. Woodson. 2018. “The Hfq Chaperone Helps the Ribosome Mature.” *EMBO Journal*. https://doi.org/10.15252/embj.201899616.

Sharma, Indra Mani, and Sarah A. Woodson. 2019. “IF3 Licenses Newly Made 30S Subunits for Translation during Stress.” *BioRxiv Biochemistry*. https://doi.org/10.1101/655696.

Soper, Sarah F Clatterbuck, Romel P Dator, Patrick A Limbach, and Sarah A Woodson. 2013. “Article In Vivo X-Ray Footprinting of Pre-30S Ribosomes Reveals Chaperone-Dependent Remodeling of Late Assembly Intermediates.” *Molecular Cell* 52 (4): 506–16. https://doi.org/10.1016/j.molcel.2013.09.020.

Stark, Martha R, and Stephen D Rader. 2014. “Spliceosomal Pre-MRNA Splicing” 1126: 137–49. https://doi.org/10.1007/978-1-62703-980-2.

Stern, Seth, Ted Powers, Li-ming Changchien, and Harry F Noller. 1989. “RNA-Protein Interactions in 30S Ribosomal Subunits : Folding and Function of 16S RRNA.” *Science* 244 (4906): 783–90.

Strader, Michael Brad, William Judson Hervey, Nina Costantino, Suwako Fujigaki, Cai Yun Chen, Ayca Akal-Strader, Chibueze A. Ihunnah, et al. 2013. “A Coordinated Proteomic Approach for Identifying Proteins That Interact with the E. Coli Ribosomal Protein S12.” *Journal of Proteome Research* 12 (3): 1289–99. https://doi.org/10.1021/pr3009435.

Sulthana, Shaheen, and Murray P. Deutscher. 2013. “Multiple Exoribonucleases Catalyze Maturation of the 3′ Terminus of 16S Ribosomal RNA (RRNA).” *Journal of Biological Chemistry* 288 (18): 12574–79. https://doi.org/10.1074/jbc.C113.459172.

Talkington, Megan W T, Gary Siuzdak, and James R. Williamson. 2005. “An Assembly

Landscape for the 30S Ribosomal Subunit." *Nature* 438 (7068): 628–32. https://doi.org/10.1038/nature04261.

Tamaru, Daichi, Kazuaki Amikura, Yoshihiro Shimizu, Knud H. Nierhaus, and Takuya Ueda. 2018. "Reconstitution of 30S Ribosomal Subunits in Vitro Using Ribosome Biogenesis Factors." *Rna* 24 (11): 1512–19. https://doi.org/10.1261/rna.065615.118.

Tan, Elliot, Timothy J. Wilson, Michelle K. Nahas, Robert M. Clegg, David M.J. Lilley, and Taekjip Ha. 2003. "A Four-Way Junction Accelerates Hairpin Ribozyme Folding via a Discrete Intermediate." *Proceedings of the National Academy of Sciences of the United States of America* 100 (16): 9308–13. https://doi.org/10.1073/pnas.1233536100.

Tomecki, Rafal, Pawel J Sikorski, and Monika Zakrzewska-placzek. 2017. "Comparison of Preribosomal RNA Processing Pathways in Yeast , Plant and Human Cells – Focus on Coordinated Action of Endo- and Exoribonucleases," 1–50. https://doi.org/10.1002/1873-3468.12682.

Torres, M, C Condon, J M Balada, C Squires, and C L Squires. 2001. "Ribosomal Protein S4 Is a Transcription Factor with Properties Remarkably Similar to NusA, a Protein Involved in Both Non-Ribosomal and Ribosomal RNA Antitermination." *The EMBO Journal* 20 (14): 3811–20. https://doi.org/10.1093/emboj/20.14.3811.

Vušurović, Nikola, Roger B. Altman, Daniel S. Terry, Ronald Micura, and Scott C. Blanchard. 2017. "Pseudoknot Formation Seeds the Twister Ribozyme Cleavage Reaction Coordinate." *Journal of the American Chemical Society* 139 (24): 8186–93. https://doi.org/10.1021/jacs.7b01549.

Weeks, Kevin M. 1997. "Protein-Facilitated RNA Folding." *Current Opinion in Structural Biology* 7 (3): 336–42. https://doi.org/10.1016/S0959-440X(97)80048-6.

Weitzmann, Carl J., Philip R. Cunningham, Kelvin Nurse, and James Ofengand. 1993. "Chemical Evidence for Domain Assembly of the Escherichia Coli 30S Ribosome." *The FASEB Journal* 7: 177–80.

Williamson, James R. 2006. "Assembly of the 30S Ribosomal Subunit" 4 (2005): 397–403. https://doi.org/10.1017/S0033583506004264.

Wilson, Timothy J., Yijin Liu, Christof Domnick, Stephanie Kath-Schorr, and David M. J. Lilley. 2016. "The Novel Chemical Mechanism of the Twister Ribozyme." *Journal of the American Chemical Society* 138 (19): 6151–62. https://doi.org/10.1021/jacs.5b11791.

Woodson, Sarah A. 2010. "Compact Intermediates in RNA Folding." *Annual Review of Biophysics* 39 (1): 61–77. https://doi.org/10.1146/annurev.biophys.093008.131334.

Woodson, Sarah A. 2011. "RNA Folding Pathways and the Self-Assembly of Ribosomes" 44 (12).

Xue, Shifeng, and Maria Barna. 2012. "Specialized Ribosomes : A New Frontier in Gene Regulation and Organismal Biology." *Nature Reviews Molecular Cell Biology* 13 (6): 355–69. https://doi.org/10.1038/nrm3359.

Yang, Zhixiu, Qiang Guo, Simon Goto, Yuling Chen, Ningning Li, Kaige Yan, Yixiao Zhang, et al. 2014. "Structural Insights into the Assembly of the 30S Ribosomal Subunit in Vivo : Functional Role of S5 and Location of the 17S RRNA Precursor Sequence." *Protein & Cell* 5 (5): 394–407. https://doi.org/10.1007/s13238-014-0044-1.

Young, R. A., and J. A. Steitz. 1978. "Complementary Sequences 1700 Nucleotides Apart from a Ribonuclease III Cleavage Site in Escherichia Coli Ribosomal Precursor RNA." *Proceedings of the National Academy of Sciences of the United States of America* 75 (8): 3593–97. https://doi.org/10.1073/pnas.75.8.3593.

Yusupov, Marat M, Gulnara Zh Yusupova, Albion Baucom, Kate Lieberman, Thomas N Earnest, J H D Cate, and Harry F Noller. 2001. “Crystal Structure of the Ribosome at 5 . 5 Å Resolution” 292 (May): 883–97.

Curriculum Vitae

Arthur Korman
IArthurKormanI@gmail.com

- Born in Ukraine
- December 13, 1990
- Moved to the United States in 1998

Education

PhD, Johns Hopkins University, Baltimore, MD 2014-Present
Program in Cell, Molecular, Developmental Biology, and Biophysics (CMDB)

Bachelor of Science, Stony Brook University, Stony Brook, NY 2009-2013
Major: Biochemistry, Summa cum laude

Research Experience

Pharmaceutical Industry Internship April 2019 – August 2019
AstraZeneca, Gaithersburg, MD
Developed analytical methods for the detection of vitamins and Polyethylenimine impurities using UPLC and HPLC systems

- Developed an LC-MS methods for the detection of five water- and fat-soluble vitamins in upstream and downstream samples
    - Using the Waters UPLC and QDa system, developed instrument and processing methods in Empower3 for vitamin identification
- Developed an Agilent HPLC method for the detection of Polyethylenimine, a contaminant
- Experience with Tecan and Microlab Nimbus instruments
- ELISA for the detection of impurities in mAbs
- Worked closely with collaborators to develop analytical methods to their specifications and time constraints

Doctoral Candidate June 2014-Present
Johns Hopkins University, Baltimore, MD
Understanding the folding of helix 1 in the 16S rRNA of the 30S small subunit of the E. coli Ribosome, and kinetics of the Twister ribozyme using single-molecule techniques

- Extensive experience in molecular biology: PCR, primer design, cloning, gel electrophoresis, site directed mutagenesis, and *in vitro* RNA transcription and purification
- Revised ribosome purification protocol, ribosomal RNA and ribosomal protein purifications, and established a 30S reconstitution protocol in the lab

- Devised and performed single-molecule microscopy experiments: using smFRET, investigated folding of the Twister ribozyme in the presence of different transition metals and crowders
- Developed a novel method for the activation and self-cleavage of Twister RNA using a photoactivated chemical group (photocage) during single-molecule microscopy
- Part of a three-lab collaboration effort to synthesize a fluorescently labelled ribozyme, containing a photocage group, and detect its folding into its secondary and tertiary structure using smFRET

Research Analyst and Undergraduate Researcher May 2012-May 2014
Stony Brook University, Stony Brook, NY
Explored the role of various effectors on the trans-translation ribosome rescue system

- Acquired extensive training and experience in molecular biology, microbiology, and biochemistry techniques: cloned numerous constructs for four different projects
- Expressed and purified DNA polymerases, optimized protocols, and assessed enzyme activity (QC) and performed biochemical and structural experiments
- Developed novel *in vivo* assays for estimation of ribosome function in pathogenic gram-negative bacteria
- Experience with flow cytometry, phage display, and *in vitro* GFP-based assays

Techniques, Software & Instrumentation

Biochemistry techniques:	**Molecular biology techniques:**	**Software:**
Protein expression and purification	PCR	MS Office
ELISA	qPCR	OpenLab
Chromatography techniques: • Affinity, ion exchange, size exclusion • ProA • GE AKTA FPLC • Agilent HPLC • Waters UPLC	Gel electrophoresis	Empower3
	Cloning	LIMS
	Mutagenesis	Unicorn
	Western and Northern Hybridization	ImageJ Graphpad Prism
	Antibody Purification	MATLAB
Water QDa Mass Detector	RNA-Ligation methods	Python
SDS-PAGE		Mathematica
Ultracentrifugation		
DNA extraction		
RNA synthesis and purification		
Fluorescence assays		
SDS-PAGE/ Native PAGE		

Publications

- Sharma, I. M., Korman, A., & Woodson, S. A. (2018). The Hfq chaperone helps the ribosome mature. The EMBO journal, 37(11), e99616.
- Korman, A., Sun, Huabing., Hua, B., Panja, S., Ha, T., Greenberg, M.M., Woodson, S.A. (in preparation)

Teaching and Leadership Experience

Johns Hopkins University, Baltimore, MD 2014-Present

- An organizer of Career Fridays committee, responsible for picking, inviting, and hosting past Johns Hopkins graduate students to talk about their current occupations
- Instructed 25-student laboratory classes in Biochemistry and Cell Biology for one semester each
- Mentored four rotation students and advised junior scientists in the lab

Stony Brook University, Stony Brook, NY

- Teaching Assistant for 2 semesters of Organic Chemistry
- Teaching Assistant for 2 semesters of Biochemistry
- Mentored other undergraduate students and multiple rotation students

Awards

- Johns Hopkins University Christian B. Anfinsen Fellowship, 2014
- Stony Brook University Eli Seifman Memorial Scholarship, 2012